DOCTOR OF PHILOSOPHY

The Link Between Fixation Location and Attention During Reading

Its Extent and Nature

Wakeford, Laura Jane

Award date:
2015

The Link Between Fixation Location and Attention During Reading: Its Extent and Nature

Laura Jane Wakeford
MA (Hons) MSc

Doctor of Philosophy
University of Dundee

April, 2015

Declaration

I declare that I am the author of this thesis and that, unless otherwise stated, all references cited have been consulted by me. The work, of which this thesis is a record, has been carried out solely by me, and has not been previously accepted for a higher degree.

Laura Jane Wakeford Date: 21/04/2015

Certificate

I certify that Laura Jane Wakeford has spent the equivalent of at least nine terms on research work under my supervision and that she has fulfilled the conditions of Ordinance No 39 so that she is qualified to submit this thesis for the degree of Doctor of Philosophy.

Dr. Wayne S. Murray Date: 22/04/2015

Acknowledgements

First and foremost, I would like to thank Dr. Wayne Murray, whose infectious enthusiasm for research has been a huge motivating and guiding force in my starting and subsequent completion of this thesis. So thank you from the bottom of my heart, my Academic Dad!

I would also like to thank my family: Mum and John, Dad and Jenny, and Nana and Dadan. In spite of the distance, your support over these past few years has been overwhelming. And to my big brother, Ben, without your encouragement I would never have contemplated attending university in the first place, so thank you for believing in me.

A very special thank you to my husband Christopher 'Doodle' for tolerating me during this process and of course, for giving us our fabulous son Archie who I hope will draw some inspiration from this journey; remember to always follow your dreams, whatever they might be.

I would also like to extend my sincere thanks to all my friends, especially: Maria Eriksson, Neil Ogg, Amy Bentley, Katja Suckow, Ross MacDonald, Glenn Williams, Clare Kirtley, Kirsty Miller, Steven Savage, Samantha Swartzman, Anue and Mike Baker-Kukona, Jens Apel, Shane and Molly Lindsay, Anne Scrymgeour and Annie Murray. This experience would not have been the same without you.

Last but not least, I would like to thank the Economic and Social Research Council, who funded both my Masters and Ph.D research.

Abstract

This thesis explores the relationship between fixation location and the locus of attention during reading. Early theories of eye movement control during reading suggested that a very tight coupling exists between the two (Just & Carpenter, 1980); however, it has since been shown that dissociations do exist. Whether these dissociations necessarily implicate parallel lexical processing, or whether they can be accommodated for within a serial-sequential framework is explored in a series of experiments.

Experiment 1 tested whether parallel lexical processing is, at the very least, psychologically plausible. Two horizontally aligned letter strings were presented simultaneously on a screen, the task being to decide whether they were physically identical or not. Even when presentation duration should have been short enough to prohibit the strictly serial processing of each word in turn, the results show clear lexical effects: high frequency words were responded to faster and with fewer errors than low frequency words. Effects of lexicality and orthography were also found. These results suggest that the two words had been processed at a lexical level in an overlapping fashion.

Experiments 2 and 3 investigated the nature and range of word n+2 preview effects. In Experiment 2, word n+1 was either a determiner or 3-letter alternative higher frequency word; in Experiment 3, word n+1 was either a 4- or a 6-letter high frequency word. A gaze contingent display change technique

was employed, where prior to passing an invisible boundary located immediately after word n, one, the other, neither or both of words n+1 and n+2 received a nonword preview. In addition to showing orthographic parafoveal-on-foveal effects stemming from word n+1, there was also evidence that word n+2 preview influenced targeting decisions on words n and n+1. Word n+2 preview effects are also found on word n+2 and in the spillover region. These effects were most wide ranging when word n+1 length was an average of 5- compared to 3-letters.

Higher-level plausibility preview effects were explored in Experiments 4-6, again using a gaze contingent display change technique. In Experiment 4 word n+1 received either an identical preview, a different but plausible one, or an anomalous, or nonword preview. Critically, an effect of plausibility arose on word n+1, with anomalous previews receiving longer inspection times than alternative plausible previews. Experiments 5 and 6 investigated the range over which these effects might occur, testing for a plausibility preview effect on word n+2. Results showed numerical, but not statistical evidence for a plausibility-related preview effect on word n+2. There were, however, clear orthographic word n+2 preview effects.

Finally, Experiment 7 experimentally tested the immediate oculomotor response to a mislocated fixation, using a text shift paradigm to simulate saccadic error and measuring the effect on lexical processing. Critically, this experiment showed that a quick error correction strategy appears to be engaged following a simulated saccadic undershoot, rather than a stay and process response. This suggests that a mislocated fixation account coupled with a stay and process response is unlikely to provide a viable explanation for lexical parafoveal-on-foveal effects.

Overall, it is suggested that current instantiations of both serial (e.g., Reichle, Warren & McConnell, 2009) and parallel (e.g., Schad & Engbert, 2012) models of eye movement control during reading appear to fail to capture major aspects of these patterns of results. The results do, however, appear to fit most parsimoniously with a perspective on eye movement control that allows for multiple words to be processed in an overlapping fashion.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Models of Eye Movement Control during Reading

If we wish to describe the processes that drive the eyes through text, we must first understand the relationship between fixation location and attention during reading. Research spanning the past forty years has combined to suggest that while the two are intrinsically connected, dissociations do exist, and it is the nature and extent of these dissociations that are the focus of this thesis.

A series of computational models have been developed to provide a framework within which certain theoretical assumptions can be tested. The primary focus of these models has been to account for the variability in the eye movement record that results from interactions between the perceptual, lexical and oculomotor systems. At the core of these models are proposals and assumptions regarding how attention is distributed during reading. They therefore provide an excellent platform on which the relationship between fixation location and attention can be explored. This chapter provides an overview of the most influential of these models, paying particular attention to how their assumptions relating to the distribution of attention have been developed in the light of emerging empirical findings.

1.1 A Tight Link Between Fixation Location and Attention: READER

The concept of tight coupling between fixation location and attention during reading can be traced back to Just and Carpenter (1980). Their conceptualisation hinged upon the immediacy and eye-mind hypotheses. According to the immediacy hypothesis, the processing of a word occurs as soon as it is encountered, with processing encompassing all levels from the initial encoding of a word through to its semantic interpretation. Their second assumption – the eye-mind assumption – suggests that readers remain fixated upon a word until that word has been processed. As an initial approximation, there appears to be good evidence that Just and Carpenter were correct. Research has consistently shown that the time spent viewing a word is related to its lexical properties. For example, there is a negative relationship between a word's normative frequency and viewing time (e.g., Just & Carpenter, 1980; Murray & Forster, 2008; Rayner & Duffy, 1986; Schilling, Rayner & Chumbley, 1998). Carpenter and Just (1983) propose that longer inspection times on disambiguating words are also indicative of a tight coupling between eye and mind. They refer to research carried out by Carpenter and Daneman (1981), which monitored participants eye movements as they read sentences like "*…some of the best bass guitarists in the country would come to this spot"*, where, upon initial inspection, the word "*bass*" could either have been interpreted as a fish or as a musical instrument. Carpenter and Daneman observed increased inspection times on the disambiguating word "guitarists" coupled with an increase in the number of regressions back to the ambiguous word "*bass*". Carpenter and Just suggested that these results indicate each word was interpreted upon initial inspection, with the above pattern of results reflecting those occasions where an incorrect initial interpretation had been made.

Thibadeau, Just and Carpenter, 1982 (see also Carpenter & Just, 1983) attempted to operationalise these assumptions in their model of the human reading system called READER. This employed the architecture of a *C*ollaborative, *A*ctivation-based, *P*roduction *S*ystem (CAPS, Newell, 1990). In this, processes are represented as condition-action productions, which are transferred to working memory once their associated activation levels reach threshold. An important feature of this model is that the activation for a certain concept can accrue from multiple processes and structures simultaneously. This allows both activated representations stored within working memory and information stored within long term memory to interact with the processes being worked upon to modify confidence for a given concept faster than if just one level had contributed at a time. Thus, the physical features of a word can interact with semantic information to influence which word out of several potential candidates becomes activated. Given that processing time in READER refers to the number of cycles required to process a word, the authors equate this 'time' to gaze duration in human readers (i.e., the summed duration of fixations on a word before a saccade is executed out of it).

Despite accounting for 79% of the variance in the gaze durations of a cohort of students reading a scientific passage (Carpenter and Just, 1983), READER has been criticised on a number of grounds. The most fundamental of these is that, as the model incorporates a total of 225 productions, it fails to achieve the succinctness and malleability of other more recent computational models. Indeed, it has been suggested that the fact that the model's productions can fire simultaneously only acts to increase the model's opaque nature (Reichle, Rayner & Pollatsek, 2003), making it somewhat redundant as a diagnostic tool for attempting to distinguish differing theories of how attention is distributed during reading.

Reichle et al (2003) also criticise Just and Carpenter's underlying assumptions from theoretical standpoint. They argue that the immediacy and eye-mind assumptions fail to capture the dynamic nature of the link between fixation location and attention. For instance, Rayner (1975) demonstrated that denying preview of a parafoveal word leads to longer fixation durations on that word when it is subsequently fixated. This undermines the immediacy hypothesis in its strictest sense, since it indicates that the processing of a word commences prior to fixation. Also, the eye-mind assumption is contradicted by the finding that effects of word difficulty can spillover onto subsequent words.

For example, frequency effects are not always confined to the word from which they originate; low frequency words are often shown to increase fixation durations on the next word in text (e.g. Rayner & Duffy, 1986). This has led researchers to adopt a more flexible conceptualisation of the relationship between fixation location and the locus of attention.

1.2. Relaxing the Link between Fixation Location and Attention: A Serial Perspective

1.2.1 Morrison's Model of Eye Movement Control during Reading

A model that assumed a somewhat looser coupling between fixation location and attention during reading is that of Morrison (1984), which combines the architectural framework of an earlier model proposed by McConkie (1979) with the principles of saccadic programming suggested by Becker and Jürgens (1979).

Following McConkie, Morrison suggested that language processing is the engine that drives the eyes through text; an attentional 'spotlight' is focussed upon just one word at a time, lexical access of that word triggers two simultaneous acts: first, it permits attention to proceed to the next word in text; second, it initiates a saccadic program to take the eye to that new locus of attention. The metaphor of attention being akin to a 'spotlight' was first popularised by Hernández-Peón (1964); however, it was Posner (1980) who demonstrated that discrete shifts in the position of this 'spotlight' could, as

Morrison (and McConkie before him) hypothesised, precede a direct eye movement to the new locus of attention.

Unlike McConkie, however, Morrison proposed that this new locus of attention has the potential to have moved beyond the word immediately to the right of the one being fixated, *providing* lexical processing on that word has also reached completion. This assumption was based on Becker and Jürgens (1979) finding that not all saccadic programs will be committed to execution. In their task, Becker and Jürgens asked participants to orient towards a target, and should that target location change, to ignore the previous location and fixate the new target as rapidly as possible. By measuring saccadic reactions to these so-called ‘double-step’ stimuli, Becker and Jürgens demonstrated that short change latencies resulted in saccades that took the eyes directly to the target’s final location, while longer latencies resulted in two saccades executed in quick succession, one taking the eye to the initial location, closely followed by another relocating the eye to the final target location. These results suggest that not only can an existing saccadic programme be cancelled by the initiation of a new one, but that saccadic programs can also run in parallel. This latter statement must be the case given that, on average, the time required to program and initiate a saccade (175-200 ms) exceeded the interval between the two saccades (minimum latency for saccades travelling in the same direction: 138ms).

Critically for Morrison's model, Becker and Jürgens established that should the initiation-to-cancellation latency be short enough – or in the case of Morrison's model, the parafoveal word be identified quickly enough - saccades could be cancelled, allowing double attention shifts to drive eye movements forward two (or potentially more) words 'downstream'. The combination of these assumptions not only allowed Morrison to provide a general account of why approximately one third of words are skipped during reading (Rayner, 2009), but also allowed him to account for the findings that short words are skipped more often than long words (Brysbaert, Drieghe, & Vitu, 2005; Rayner & McConkie, 1976; Rayner, Sereno & Raney, 1996; Rayner, Slattery, Drieghe & Liversedge, 2011), and high frequency words are skipped more frequently than low frequency words (Henderson & Ferreira, 1993; O'Regan, 1979; Rayner & Fischer, 1996; Rayner, Sereno & Raney, 1996), as these words will have a higher probability of being identified parafoveally before their saccadic program is committed to action. Morrison's model therefore suggests an intermittent separation between fixation location and attention during reading, with a closely coupled 'cat and mouse' relationship that allows attention to precede an eye movement, before the eye 'catches up'.

Indeed, Morrison's model is able to account for a variety of phenomena in the eye movement record. Importantly, it provides a plausible explanation for how the lexical properties of a word can be implicated in driving the eyes through text despite the majority of each fixation apparently being dedicated to the post lexical process of planning a saccade (average fixation duration 225ms to 250ms, Rayner, 2009; average saccadic latency: 175ms to 200ms; Berker & Jürgens, 1979). By postulating that words receive a parafoveal preview benefit equivalent in duration to the saccadic latency prior to fixation, Morrison's model is able to circumvent this apparent temporal contradiction.

Evidence indicating that readers can obtain a parafoveal 'preview benefit' comes from research manipulating parafoveal preview. For example, in the 'moving window technique' (McConkie & Rayner, 1975) the region of available text moves with the participant's gaze as they read. Outwith this region, letters are replaced by either X's or alternative letter-strings. The effective span of apprehension is determined by assessing the limits to which reading can progress normally before the window size affects reading rate. From these experiments, it has been found that for alphabetic scripts, such as English, readers obtain parafoveal information from 3-4 character spaces to the left, and up to 14-15 character spaces to the right of fixation. Rayner, Well, Pollatsek and Bertera (1982) went on to show that reading rate diminished by only 10% if the foveal word and the word to the right of fixation were always available, and when the window was extended to include two words to the right of fixation, reading rate only deviated minimally from the norm. This suggests that readers' span of attention is not localised to the currently fixated word. Rather there is an asymmetrical span of apprehension encompassing up to 20 character spaces.

Another gaze contingent technique coined the 'boundary paradigm' (Rayner, 1975), enables researchers to manipulate the preview of specific target words prior to fixation. Typically, a boundary is located immediately after the pre-target word and it is only once the eye passes this invisible boundary that the correct preview of the word appears. Since the change is triggered during a saccade, when the visual system is suppressed, readers are typically unaware of the word change occurring (Matin, 1974). Nevertheless, it is usually found that fixations following an identical preview are shorter than if parafoveal preview had been denied, indicating that words typically undergo some degree of processing prior to receiving a direct fixation. In a meta-analysis of studies employing the boundary paradigm, Hyönä, Bertram and Pollatsek (2004) report that, on average, gaze duration increases by between 30ms and 50ms following an invalid preview.

Morrison's model therefore provides an elegant account of (a) how cognitive processing can influence fixation durations during reading despite the tight time constraints imposed by saccadic latency, and (b) why some word types tend to be skipped more frequently than others. The moving window technique and boundary paradigm have proved to be important tools in validating Morrison's assumption that a shift in attention precedes the movement of the eye.

This degree of de-coupling between fixation location and attention does not, however, afford an explanation for spillover effects. As described above, the processing difficulty of a word can also modulate fixation duration on subsequent words (e.g., Rayner & Duffy, 1986). Given that the model suggests that the completion of lexical processing of a word is the trigger for both a shift of attention and for the programming of a saccade to the next word, variables effecting lexical difficulty should exhibit effects confined to fixations on that word only and never spillover. Consequently, in order to account for factors such as frequency effects spilling over, proponents of Morrison's model suggested that these were likely to arise as a consequence of increased difficulty integrating low frequency words into the text (Rayner, Sereno, Morris, Schmauder & Clifton, 1989).

Another finding that Morrison's model cannot readily account for is the modulation of parafoveal preview benefit by foveal difficulty. Henderson and Ferreira (1990) demonstrated this effect by simultaneously manipulating parafoveal preview of a target word and pre-target word difficulty. They found that when the pre-target word was difficult (low as opposed to high frequency in Experiment 1 and increased syntactic complexity in Experiment 2), a significant reduction in target word preview benefit was observed. Again, this finding cannot be reconciled with a model of eye movement control that ties the decision to plan a saccade out of a word with the moment when lexical access of that word is completed.

1.2.2. The E-Z Reader Model of Eye Movement Control During Reading

Despite the inadequacies of Morrison's Model in explaining some effects, it has laid the foundations for what has become one of the most influential models of eye movement control during reading: E-Z Reader. Its primary architect, Reichle, describes their model as "a family of models developed over the past decade to provide increasingly sophisticated descriptions of how various perceptual, cognitive, and motor processes guide readers' eye movements" (Reichle, 2011, p771). An important aspect of this model is that it is modular (in the sense suggested by Fodor, 1983). While early versions of the model included just two distinct modules relating to word recognition and the execution and programming of saccades, it could be later extended with relative ease to include modules that account for landing site distributions (E-Z Reader 6; Reichle, Rayner & Pollatsek, 1999) and post-lexical processing effects (E-Z Reader 10; Reichle, Warren & McConnell, 2009).

An important functional difference between the E-Z Reader model and its predecessor is that E-Z Reader is a quantitative model that allows certain theoretical assumptions to be 'tested' on a large number of statistical subjects. If the findings from experimental data can be modelled using E-Z Reader, the assumptions specified in E-Z Reader are considered to be verified, thereby reflecting the potential underlying processes involved in the progression of reader's eye movements through text. This clearly differs from Morrison's

Model, which due to its qualitative nature, was only capable of providing verbal descriptions and hypotheses relating to eye movement control during reading.

The E-Z Reader model's transparent nature combined with its ability to make quantitative predictions has made it the model of choice for testing the possibility that it is lexical processing that drives the eyes through text, with attention only being focussed upon one word at a time. Given the importance of this model in testing the link between fixation location and attention, a full description will now be presented.

The E-Z Reader model retains several of Morrison's key assumptions. For example, it adopts a strictly serial sequential architecture, in which only one word can be lexically processed at a time and with lexical access on that word permitting attention to proceed to the next word in text. E-Z Reader also incorporates the two-stage saccadic programming mechanism that allows for the parallel programming of saccades (Becker and Jürgens, 1979). However, in a fundamental extension of Morrison's model, Reichle, Pollatsek, Fisher & Rayner (1998) decoupled attention shifts from the decision to programme a saccade, allowing each to be triggered by the completion of a different stage of lexical processing. This decoupling has remained a central assumption throughout all subsequent versions of the model.

In the model, the word recognition process comprises two discrete stages, termed L1 and L2. Completion of the first stage, L1, signals that lexical access is imminent and triggers a saccade to be planned to the next word in text; while completion of the second stage, L2, indicates that full lexical access has occurred and permits attention to shift to the next word in text. The time required to complete each stage of lexical access is determined by the word's normative frequency (as tabulated by Francis & Kučera, 1982) and predictability (determined by cloze task). The time required to complete L2 in the model has always been treated as a fixed proportion of the total time taken to complete L1 prior to any adjustments of L1 (see below), with a value of .5 in Version 10 of the model.

Early versions (version 1-7) of the E-Z Reader model assumed a multiplicative relationship between the effects of word frequency and predictability, with predictability having its biggest influence on processing times when bottom-up processing is slow. However, Rayner, Ashby Pollatsek and Reichle (2004) conducted an experiment in which they orthogonally manipulated a critical word's normative frequency and predictability and found that - for durational measures at least - only a weak (non significant) relationship existed. E-Z Reader was therefore adapted to assume an additive relationship between these two variables. While the architects acknowledged in their first publication of the model (Reichle et al, 1998) that many additional factors will likely influence the durations of L1 and L2 (e.g., recency of usage,

age of acquisition and how many neighbours the word has), additional factors influencing processing time have remained absent from all subsequent versions of the model.

Decoupling saccadic programming and attention shifts has allowed the E-Z Reader model to explain how foveal difficulty can modulate preview benefit within a serial architecture. Specifically, because the time required to programme and execute a saccade is always independent of a word's difficulty, upon completion of L1, the saccade will execute with the same latency for both easy and difficult to process words. However, because the time required to complete L2 is a function of the word's difficulty, L2 will complete, and therefore attention will shift to the next word in text, sooner for easy-to-process compared to difficult-to-process words. The word's difficulty will therefore modulate the amount of parafoveal preview benefit that a word to the right of fixation can accrue. Spillover effects can be explained by this mechanism: difficult to process foveal words reduce the amount of time available to parafoveally process a word to the right of fixation before a saccade is executed to that word, thereby inflating fixation durations following a difficult to process word. Decoupling attention shifts from the decision to plan a saccade has therefore resulted in a more dynamic relationship between fixation location and attention during reading, with the degree to which the two are dissociated being a function of the fixated word's difficulty.

Pollatsek, Reichle and Rayner (2006, p10) emphasise that defining L1 and L2 as sequential stages was, in part, a modelling convenience and that they “do not necessarily conceptualise them that way”. Reichle, Pollatsek and Rayner (2006) suggest three different ways in which the distinction between these two stages can be envisaged. First, the two stages might reflect temporal differences in the times that different codes take to become activated; for instance, the orthographic code might become available before phonological and semantic codes. Alternatively, the distinction could reflect L1 being equated with a rapidly accessible recognition process, while L2 reflects a slower retrieval process (Atkinson & Juola, 1973, 1974; Yonelinas, 2002; cited in Reichle et al, 2006). Finally, the two stages could reflect pre- and post-lexical access. Throughout the ten versions of E-Z Reader, its architects have remained agnostic regarding the precise relationship between these two stages. Rather they simply maintain the notion that L1 reflects the processing system - based on past experience - reaching some threshold that indicates lexical access is imminent, and this allows the saccadic programming system to get a ‘head start’.

Regardless of the conceptualisations of L1 and L2 that one wishes to adopt, the processes involved in L1 will - by nature of its position in the word recognition process - always be associated with the early stages of processing, and will therefore be tied to the extraction of visual features from the page. However, according to the E-Z Reader model, there are times when this process may be redundant as rapid top-down processing may intervene to supply the processing system with sufficient information to “fill in the gaps” in the sentence’s meaning (Pollatsek et al 2006, p12). This assumption is instantiated within the model by allowing a word’s predictability to fully affect L1, such that the time necessary for this stage to complete can be reduced to zero msec for a word that is highly constrained by its preceding context. Thus, in the event of an attention shift to a highly predictable parafoveal word, a new saccadic programme will be programmed immediately, and this can override any ensuing program to fixate that word, causing it to be skipped. This assumption elegantly explains the finding that highly predictable words are more commonly the recipients of a skip than are words of low predictability (e.g., Balota, Pollatsek & Rayner, 1985; Drieghe, Brysbaert, Desmet & De Baecke, 2004; Rayner, Binder, Ashby, & Pollatsek, 2001; Rayner & Well, 1996). Given that time is always required to incorporate the meaning of a word within its sentence frame, L2 is never attenuated in the model (Reichle et al, 2011).

Another factor that influences the first stage of lexical processing in the model is visual acuity. Resolution is at its highest within the fovea, but our ability to extract fine grain detail rapidly decreases with increasing eccentricity into parafoveal and peripheral regions (Chung, Mansfield & Legge, 1998; Legge, Mansfield, & Chung, 2001). Rayner and Morrison (1981, see also Miellet, O’Donnell and Sereno, 2009) have shown that the primary determinant of acuity during reading should be measured in character spaces rather than degrees of a visual angle. They found that long words or words that are far from the fovea took longer to process than short words or words that fall close to the fovea. The reason for this is likely routed in a trade-off between visual acuity and letter size: while larger letters occupy a larger visual angle and will thus be subject to a greater loss of visual acuity, this is offset by the lower level of acuity required to discriminate larger letters.

To capture the impact acuity has upon fixation durations during reading, the E-Z Reader model (version 6 and onwards) allows visual acuity to modulate the time required to complete L1. This is achieved by incrementally slowing the processing rate for each letter the further that letter falls from the point of fixation. This explains word length effects in which short words receive shorter fixation durations compared to long words (e.g., Just & Carpenter, 1980; Pollatsek, Juhasz, Reichle, Machacek & Rayner, 2008; Rayner & McConkie, 1976), because the peripheral letters in long words are further from the centre of vision and therefore take longer to encode. This also suggests that parafoveal words will be encoded at a slower rate than fixated words (Rayner & Morrison, 1981); again due to the disparity between the point of fixation and the number of letters intervening before those parafoveal words.

The second stage of word processing, L2, is unaffected by visual acuity constraints, as visual acuity should only influence feature extraction and therefore only the early stages of word processing. Indeed, this assumption is consistent with Reingold and Rayner's (2006) finding that visual degradation only increased fixation durations on the word being fixated, and did not influence spillover effects, since these will be modulated by the duration of L2.

Thus far the link between fixation location and attention within the E-Z Reader model is only marginally weaker than the ideas proposed by Just and Carpenter (1980). For the most part, fixation location and attention are assumed to overlap, with dissociations only occurring when there is a mismatch between the time required to plan and execute a saccade and the time required to complete L2; at which point either attention runs ahead of the eye (for easy to process foveal words), or occasionally, the eye runs ahead of attention (for an especially difficult to process foveal word). But this cat and mouse relationship between attention and eye position only permits attention to proceed to the subsequent word once the currently fixated word has received full lexical access.

On the basis of this architecture, therefore, the difficulty of a parafoveal word should never modulate fixation duration on the currently-fixated word. Such 'parafoveal-on-foveal' effects have, however, been reported in the literature, most notably in cases where there is a peculiar orthography in the parafoveal word position. This can be seen in studies using the gaze contingent display change paradigms described above, in which an orthographically illegal preview of a parafoveal word influences fixation durations on the currently fixated word (e.g., Angele & Rayner, 2011; Hyönä & Bertram, 2004; Inhoff Starr & Schindler, 2000; Starr & Inhoff, 2004). Such effects are, however, variable, often only arising when fixations fall on the final three character spaces of a foveal word (e.g., Drieghe, Rayner and Pollatsek, 2008; Rayner, 1975) and they often fail to be observed at all (Rayner, Juhasz & Brown, 2007; White & Liversedge, 2004). Despite such inconsistency in the appearance of orthographic parafoveal-on-foveal effects across experiments, recent reviews of the literature have generally converged in acknowledging their existence (Radach & Kennedy, 2004; 2013; Schotter, Angele & Rayner, 2012).

The E-Z Reader model has evolved to account for these effects by including an early pre-attentive stage of visual processing. According to Reichle et al (2003), this pre-attentive stage extracts featural information from a written passage of text prior to focussed attention converting those features into perceptual wholes; indeed, such a distinction is not without precedence in the literature (e.g, Pollatsek & Digman, 1977; Treisman & Galade, 1980). This pre-attentive stage is thought to extract low-spatial frequency information from a page in parallel, which can be later utilised by the oculomotor system for determining the targets of upcoming saccades. The time required to transfer this pre-attentive information from the retina to the visual cortex in the most recent versions of the E-Z Reader model (versions 9; Pollatsek et al, 2006; and 10; Reichle et al, 2009) is 50-ms, which reflects findings of neurological research (e.g., Foxe & Simpson, 2002; Mouchetant-Rostaing, Giard, Bentin, Aguera, & Pernier, 2000; Van Rullen & Thorpe, 2001). Visual acuity constraints restrict the uptake of visual information at the pre-attentive stage in the same way as for the first stage of lexical processing. Despite such constraints, Reichle et al (2003) suggest that the pre-attentive stage is capable of detecting orthographic irregularities in the parafovea, as these unusual feature combinations "pop out" from the page (Reichle et al, 2006; p8), driving the above-mentioned orthographic parafoveal-on-foveal effects. The addition of this pre-attentive stage allows the E-Z Reader model to maintain the suggestion that focussed attention only moves to the next word in text once the currently fixated word has been fully identified, thereby allowing it to retain its strictly serial-sequential assumption.

Up to this point, discussion has centred on what Reichle et al (2006) describe as being the 'front end' of their model. However, the E-Z Reader model also makes precise predictions relating to the programming and execution of saccades, derived from the oculomotor module of the model. Following Morrison, saccades in the E-Z Reader model are programmed in two stages: M1 and M2, the former of which can be cancelled, while during the latter the saccade is committed to execution. If a replacement saccade is programmed during M1, the original saccadic program will be cancelled resulting in a word skip; however, if the new saccade is initiated during M2, the original saccade will be committed to execution and the two saccadic programs will run in parallel.

The first labile stage, M1, is further divided into two stages, a preparatory stage, which determines the upcoming target, and a translation stage, which determines distance to that target and therefore the force required to execute the saccade. During the non-labile stage, M2, the saccade cannot be cancelled according to the model, owing to the command having been sent to the brainstem circuitry that carries out the movement. As implemented in the E-Z Reader model, M1 and M2 take, on average, 100ms and 25ms to complete, while the actual execution of the saccade is estimated to take 25ms. This aspect of the model is simply an extension and formalisation of the assumptions proposed in Morrison's model, which were inspired by Becker and Jürgens (1979).

E-Z Reader 6 (Reichle et al, 1999) did, however, extend Morrison's model to account for landing site distributions on words. This addition was spurred by the criticisms of Reilly and O'Regan (1998), who suggested that the E-Z Reader model would struggle to account for landing site distributions, an assumption based on their own failure to simulate these effects using the architectural framework of Morrison's model. To counter this criticism, Reichle et al (1999) expanded their oculomotor module to include landing site distributions. Before outlining these additional assumptions, it is necessary to briefly consider the effects Reichle et al needed to model.

Landing site distributions on words are typically broad and assume a normal distribution with the peak offset to the left of word centre, and with the tails stopping abruptly at word boundaries. The peak of this inverted U-shape function is often referred to as the "*preferred viewing location*" (Rayner, 1979). Indeed, it is not surprising that readers should aim for a word's midpoint, given that when presented in isolation, words are identified most rapidly from a central position, referred to as the "*optimal viewing location*" (O'Regan, 1981; O'Regan & Jacobs, 1992; O'Regan, Lévy-Schoen, Pynte & Brugaillère, 1984). It is surprising, therefore, that during natural reading, these central positions are associated with the longest initial fixation durations, and these incrementally decrease with increasing eccentricity from the optimal viewing position. This has become known as the "*inverted optimal viewing position*", or IOVP effect (Vitu, McConkie, Kerr, & O'Regan, 2001). Nuthmann, Engbert and Kliegl (2005; 2007; Engbert, Nuthmann & Kliegl, 2007) have suggested an explanation for these counterintuitive findings: while the optimal viewing position may indeed be the target for an upcoming fixation, error in the oculomotor system could result in saccades either under- or over-shooting their intended target location. This is likely to trigger an error-correcting strategy designed to relocated the eye to a more optimal position. If one assumes that there is an increasing probability of initiating a quick corrective saccade in response to an erroneous fixation the further that fixation falls from a word's centre, it seems entirely plausible that this tendency could give rise to the IOVP effect.

Research conducted by McConkie, Kerr, Reddix, and Zola (1988) suggests that there is indeed error in the oculomotor system. McConkie et al plotted within-word initial landing site distributions and found that, for saccades travelling 6 to 7 characters, landing site distributions assumed a normal distribution with a peak located at a word's centre. This peak shifted, however, by approximately half a character space for every character the saccade deviated from 7 characters, with near and far launch sites being associated with shifts to the right and left, respectively. They also noted a reduction in this shift following increased fixation durations at the launch site. McConkie et al interpreted these results as reflecting a preference to target a word's centre, but with a systematic range error that can cause saccades to under- or overshoot their targets. This systematic range error, they suggest, can be reduced following longer prior fixations, since this is likely to give rise to more accurate targeting. According to McConkie et al, this error is most likely related to the range error found in motor movements generally, and ocular muscle control specifically, in which the eyes tend to overshoot close targets and undershoot more remote ones (e.g., Kapoula, 1985). They also implicate a random error component in the distributions. This error, they suggest, gives rise to less accurate targeting following more remote launch sites with these associated with increased variability in landing site distributions. The application of these principles will now be described in the context of the E-Z Reader model, since as the architects of the model acknowledge, these suggestions were incorporated “more or less directly” from McConkie et al’s data and analyses (Reichle et al, 2003, p454).

Since their inclusion in the model (model 6 onwards), the assumptions regarding landing site distributions and the main principles for explaining them have remained relatively static. The model assumes that readers target the OVP. However, the actual saccadic length is the sum of three components: (a) the intended saccade length, (b) systematic range error, and (c) random error. As per McConkie et al, the preferred saccade length in the E-Z Reader model is specified as being 7 character spaces; every character space a saccade’s length deviates from 7 characters, incurs a systematic error of approximately half a character space. The magnitude of this error is however scaled by a linear function of the natural logarithm of launch site fixation duration, which acts to reduce the magnitude of error as launch site fixation duration increases. Random error is simply represented in the model as a normally distributed Gaussian function, with a mean equal to zero and a standard deviation that increases with the length of the programmed saccade. After running a simulation on the Schilling et al Corpus using 4-7 character words, Reichle et al (1999) concluded that the E-Z Reader model did “a fairly good job” (p 4407) at replicating first pass initial fixations positions on a word.

To answer the question of how well the E-Z Reader model accounts for the IOVP effect reported by Vitu et al (2001), we must first look at how the model accounts for refixations. The intra-word distribution of refixations follows an asymmetrical U-shape function based on landing site position, with a greater number of refixations to the left compared to the right extremities of a word, and with the fewest refixations triggered from central word positions (Rayner, Sereno & Raney, 1996).

Reichle et al (1999) first attempted to model these effects in E-Z Reader 6. They assumed that a saccadic plan to refixate a word is initiated automatically upon fixation; as with the programming of inter-word saccades, these intra-word saccadic programs comprise two stages: M1 and M2. Should the refixation program enter M2, the refixation will be committed to action. However, should L1 on the fixated word complete first, a new saccadic program to the next word in text will override the refixation. As the time required to complete L1 in the model is a function of word frequency, the model correctly suggests that fewer refixations will occur on high compared to low frequency words (e.g. Rayner et al, 1996). Also, given the acuity constraints discussed previously, this conceptualisation accurately captures the asymmetrical distribution of refixations. In addition, long words and/or fixations falling at the ends of words, will – due to acuity constraints - require more time to complete L1 than when short words and/or more central locations are fixated, therefore, allowing the M1 stage of the refixation programme to complete first. E-Z Reader 6 therefore predicts more rapid error-correcting refixations from a word's extremities, allowing it to account for the IOVP effect.

However, despite the fact that E-Z Reader 6 was able to correctly model the distribution of refixations, it produced first fixation durations on low frequency words that were unreasonably short. This came about because in order for the correct proportion of refixations to be estimated, the M1 stage for these refixations had to be set low enough to allow them to complete before the first stage of lexical processing. The architects of E-Z Reader 7 (Reichle et al, 2003) attempted to remedy these flaws by suggesting that refixations are not initiated automatically upon fixation; rather they are initiated with a probability determined by word length. However, these ideas were abandoned with the introduction of E-Z Reader 9 (Reichle et al 2006).

In E-Z Reader versions 9 and 10, the probability of initiating a rapid corrective refixation is proportional to the absolute distance between where the initial fixation falls within a word and the OVP. In these newer versions, a refixation cannot be programmed within the first 50ms of fixation; the theory being that visual feedback is required before a decision can be made as to whether to relocate to a better location. Two final stipulations prevent a refixation from overriding an inter-word saccade: a refixation cannot be planned if (a) there is an existing labile program already in operation or if (b) a saccade is in execution. These new assumptions allow the model to predict increasing numbers of rapid error-correcting refixations the further the first landing position falls from the OVP, and to allow longer words to receive more refixations, as they have an increased probability of receiving initial fixations that fall farther from the OVP (Vergilino & Beauvillain, 2000). These assumptions are sufficient to enable the model to account for the IOVP effect given that quick refixations will be initiated with increasing eccentricity from the OVP, giving rise to shorter initial fixation durations in these peripheral regions of the word. In summary, the mechanisms responsible for the IOVP effect in the E-Z Reader model align with Nuthmann et al's (2005; 2007) suggestion that fixations falling in erroneous locations might give rise to quick error-correcting saccades. However, they do deviate from Nuthmann et al in that the error-correction strategy arises from within-word targeting error and not saccades that miss their target words.

So-called 'mislocated fixations' (i.e., saccades that land on the wrong word) do, however, feature heavily, if only verbally and not quantitatively, in current literature related to the E-Z Reader model. McConkie at al (1988) were the first to propose that error in the oculomotor system might cause a saccade to either under- or over-shoot its intended target to the extent that a fixation will occur on adjacent words. This suggestion has been popularised in recent years (e.g., Drieghe, Rayner & Pollatsek, 2008; Nuthmann et al, 2005; 2007) and provides a plausible explanation for why there is truncation of the tails of the above mentioned landing site distributions. Engbert et al (2007) estimated the extent to which mislocated fixations might occur during normal reading. Their simulations suggested that over- and under-shoots might account for as many as 16% and 7.2% of all fixations on words, respectively. Given the size of these proportions, any systematic response to such errors could have a significant impact on the overall eye movement record and they are therefore of concern to models of eye movement control, such as E-Z Reader.

In contrast to the quick error-correction strategy proposed by Nuthmann et al (2005), the proponents of the E-Z Reader model have advanced an alternative explanation for how the oculomotor system might respond to mislocated fixations. Specifically, Rayner, Warren, Juhasz and Liversedge (2004) suggest that rather than initiating a quick error-correcting saccade, the reader might adopt an alternative ‘stay and process’ strategy. Such a strategy allows the model an account for how the lexical properties of a word to the right of fixation could modulate inspection time despite the model’s serial sequential framework apparently precluding the occurrence of such effects. This type of account was further popularised by Drieghe et al (2008) who suggested they had obtained evidence for the stay and process strategy following saccadic undershoots. This research and the associated issues will be discussed more fully in Chapter 6.

Thus, in addition to accounting for orthographic parafoveal-on-foveal effects (via a low-level attentional scan), proponents of the E-Z Reader suggest that their model is now capable of accounting for lexical (e.g., Hyönä & Bertram, 2004; Kennedy, 1998; 2000; Kennedy, Pynte & Ducrot, 2002; Klielg, Nuthmann & Engbert, 2006) and higher level (e.g., Inhoff, Radach, Starr & Greenberg, 2000; Murray & Rowan, 1998; Rayner, Warren Juhasz & Liversedge, 2004) parafoveal-on-foveal effects. Unfortunately, to date, neither the low level attentional scan nor the stay and process response to mislocated fixations has been instantiated in the model. Therefore arguments regarding the model's ability to simulate such effects remain largely verbal rather than quantitatively determined.

It should be noted that despite the possibility that this explanation can account for lexical parafoveal-on-foveal effects within a serial framework, there are also aspects of it that appear to be incongruent with some of the central mechanisms of the E-Z Reader model. First, the acuity constraints that are built into the E-Z Reader model (version 6 onwards) would appear to mitigate against such a stay and process strategy. Second, there is no apparent reason why the refixation mechanism (as determined by foveal distance from the OVP) should not extend to fixations that fall an extra one or two character spaces away from the OVP and land on an adjacent word; this should produce a quick error-correcting response and not a stay and process response to a mislocated fixation.

It should also be noted that for mislocated fixations to account for the IOVP effect as suggested by Nuthmann et al, 2005; 2007, a quick error-correcting saccade needs to be initiated, while to account for lexical parafoveal-on-foveal effects as suggested by Drieghe et al (2008), a stay and process response must be initiated. As Engbert and Kliegl (2011) highlight, these two positions must be mutually exclusive. This final point and an investigation of the relative merits of these assumptions will be returned to in Chapter 6. However, in the context of the present discussion, such an interpretation of apparent lexical parafoveal-on-foveal effects also has repercussions for how the E-Z Reader model instantiates the link between fixation location and attention.

Prior to the adoption of a mechanism involving mislocated fixations coupled with a ‘stay and process’ strategy as a characteristic of the model, the link between fixation location and attention has been hypothesised to remain relatively tight. Indeed, it only really deviated from Just and Carpenter’s (1980) immediacy and eye-mind hypotheses to allow attention to (briefly) precede an eye movement, or – less typically, following a ‘difficult’ to process word – for the eye to (briefly) move ahead of attention.

Decoupling attention shifts from the decision to plan a saccade has allowed the E-Z Reader model to accommodate a variety of effects without compromising its key assumption of a tight coupling between fixation location and attention. Once the concept of a mislocated fixation is factored into the model, however, this decoupling is stretched far beyond that proposed in the original model introduced in 1998. Previously, while the two were often dissociated, one was always en route to the location of the other. Now, the proponents of the model actively promote the possibility that this ‘pull’ to be together can breakdown – allowing the two to become dissociated.

The model described so far resembles what its architects envisage happens when perceptual, oculomotor and lexical processes interact to drive the eyes through text and no higher level problems intervene; that is, it is the default reading state. However, E-Z Reader 10 (Reichle et al, 2009) extends previous versions of the E-Z Reader model to account for how post-lexical integration might interact with this default state, potentially causing disruption. In the model, the term “post lexical integration” encompasses all post lexical activities, ranging from placing a word into a syntactic structure, through to incorporating its meaning into a discourse model. The inclusion of an integrative stage is specified in the model as follows: The mean time required to complete the integration stage in the model is 25ms, and is initiated as soon as a word has been lexically identified (that is, L2 has completed). The architects acknowledge that 25ms might appear short, however, they argue that it is sufficient to provide a “good enough” level of integration (e.g., Ferreira, Bailey & Ferraro, 2002; Ferreira & Patson, 2007; Swets, Desmet, Clifton & Ferriera, 2008), such that the forward progression of lexical processing need not be interrupted. Successful integration then permits predictability information to influence lexical processing of the next word in text. Integration fails either (a) immediately after lexical identification of a word, or (b) when the word cannot be integrated before the next word in text has been lexically identified. Following integrative failure, both attention and the eyes regress to either the word causing the difficulty, or to an earlier word – in E-Z Reader 10 this is always (due to modelling convenience) the word immediately to the left of the one causing the difficulty. Finally, it is assumed that regressive saccades require an extra 30ms to programme in the M1 stage than progressive saccades.

Taken together, these assumptions allow the model to either halt the progression of the eyes through text, or to cause both eye movements and attention to be directed back to the region of difficulty, or earlier. The addition of these assumptions therefore provides some explanation for the rapid influences of syntactic ambiguities and semantic violations in the eye movement record (e.g., Frazier & Rayner, 1982; Rayner et al 2004; Warren & McConnell, 2007). The addition of the post lexical integration module in E-Z Reader 10 does not therefore increase the decoupling between fixation location and attention beyond that found in earlier versions of the model. Indeed, the response is quite to the contrary, with a regrouping of fixation location and attention following integration failure.

In summary, the E-Z Reader model provides a transparent and parsimonious account of how fixation location and attention are linked within a serial framework of eye movement control during reading. The model proposes a close relationship between the two, with attention only being free to move forward once the currently fixated word has received full lexical access. On the surface, it would appear that the inclusion of a pre-attentive visual processing stage allows the model to account for orthographic parafoveal-on-foveal effects. While this seems a plausible addition based on previous research, it has not yet been confirmed whether the E-Z Reader model is capable of simulating these effects with the inclusion of such a mechanism - and whether doing so might be at the expense of simulating other benchmark effects that it can currently simulate. Whether the stay and process response to a mislocated fixation can provide a plausible explanation for lexical parafoveal-on-foveal effects seems somewhat more tenuous. What is clear, however, with the inclusion of the pre-attentive stage of visual processing and the suggestion that a word can be processed from an erroneous fixation, is that the E-Z Reader model has adapted in such a way that makes it increasingly difficult to distinguish – at least experimentally – from models that assume that lexical processing of multiple words can occur in parallel during reading. Examples of these will be discussed next.

1.3 Relaxing the Link Between Fixation Location and Attention: A Parallel Perspective

1.3.1. Processing Gradient Models:

Processing Gradient (PG) models, such as SWIFT (Engbert, Longtin & Kliegl, 2002; Engbert, Nuthmann, Richter & Kliegl, 2005; Richter, Engbert & Kliegl, 2006; Schad & Engbert, 2012) and Glenmore (Reilly & Radach, 2006) propose a somewhat looser link between fixation location and the focus of attention during reading. While the architects of PG models agree that cognitive factors can influence the progression of eye movements through text, they also suggest that a model proposing serial sequential attention shifts is too restrictive to provide a viable account for a variety of phenomena observed in the eye movement record, such as the above-mentioned lexical parafoveal-on-foveal effects (Engbert & Kliegl, 2011). Instead, these models assume that multiple words falling within a gradient of attention can be lexically processed in parallel.

To be clear, while E-Z Reader allows multiple words to be lexically processed within one fixation (due to attention shifts), a parafoveal word can only be processed once the currently-fixated word has received full lexical access. PG models on the other hand suggest that all words within the gradient of attention can undergo lexical processing simultaneously.

One PG model - SWIFT - was designed and implemented in a fully quantitative manner to be a viable alternative to models, such as the E-Z Reader model, that restrict lexical processing to just one word at a time. Although not the only model in its class, the SWIFT model is the most developed and extensively tested model of this variety. As such, the following section will concentrate on the theoretical underpinnings and associated mechanisms of this model.

1.3.2 The SWIFT Model of Eye Movement Control during Reading

As discussed above, experiments employing the moving window technique suggest that a skilled reader's effective span of apprehension during reading is asymmetrical, encompassing roughly 3-4 character spaces to the left and up to 14-15 character spaces to the right of fixation (McConkie and Rayner, 1975; 1976). This breadth of and asymmetrical distribution of attention was incorporated within the first version of SWIFT (Engbert et al, 2002), which allowed the simultaneous lexical processing of up to four words. The gradient of attention encompassed the foveal word, one word to the left and two words to the right of fixation. In this early version, processing rate is determined by eccentricity at the word level with the foveal word (word n) receiving the fastest rate of processing, while its two spatially adjacent neighbours (word n-1 and word n+1) receive lesser amounts and the word two to the right (word n+2) receives the slowest rate of processing.

While this first version neglected word length, its successor, SWIFT II (Engbert et al, 2005) accounted for the distribution of attention and acuity constraints at the letter rather than word level. The gradient of attention can be conceptualised as an asymmetric Gaussian distribution, with the gradient of the function being determined by two free parameters, one for the steep decline in processing efficacy to the left of fixation and another for the comparatively shallower reduction in processing efficacy to the right of fixation. According to this implementation, the processing rates for all the individual letters contribute toward a word's overall processing rate. With this mechanism, SWIFT II is therefore able to account for word length effects during reading (e.g., Just & Carpenter, 1980; Pollatsek et al, 2008; Rayner & McConkie, 1976). It also provides a more realistic conceptualisation than their first model of the limits of the effective span of apprehension during reading. Using SWIFT to run simulations on the Potsdam Corpus it was evident that, typically, just 3-4 words were lexically processed simultaneously; only extending to five words on 5% of occasions (Richter et al, 2006). All following discussion will be based on this second version of SWIFT unless otherwise stated.

SWIFT adopts the theoretical assumptions of the dynamic field theory of eye movement preparation (Erlhagen & Schoner, 2002). as the foundation for the interaction between perceptual, lexical and oculomotor systems. In SWIFT, the dynamic field can be conceptualised as a one-dimensional saliency map, in which words falling on a predetermined horizontal axis receive varying degrees of activation. The relative activation levels of words are influenced by the eccentricity constraints discussed above, such that activation levels are adjusted more rapidly for input closer to the point of fixation. Activation levels are also a function of a word's normative frequency and its predictability; the exact processes by which these exert their influences are discussed below. Words falling outwith the horizon of attention generally receive negligible levels of activation. The relative activation levels evolve dynamically as a function of time and provide the oculomotor system with potential targets for upcoming saccades.

According to SWIFT, a word holding the highest level of activation at a specific time is determined as the target, whether that target results in a progressive or regressive saccade, or whether it causes a refixation on the currently fixated word, or a skip of the word to the right of fixation. Target selection can therefore be conceptualised as a competitive process among all activated words within the effective span of apprehension. The architects of SWIFT (Engbert et al, 2005) argue that imposing one common mechanism responsible for all types of saccade provides their model with an increased degree of parsimony and is one of the model's key strengths over its main rival – the E-Z Reader model – that requires extra mechanisms and assumptions to account for regressions and refixations (as discussed above).

In contrast to the assumptions of the E-Z Reader model, the primary determinant of saccadic timing is not lexical processing. The architects of the SWIFT model assume that the processor is designed to progress through text maintaining a mean rate of eye movements, the intervals of which are determined by a reader's preferred reading rate and the difficulty of the text being read (Richter et al, 2006). This system is autonomous and is what the architects refer to as a "dumb" default strategy, since, as Deubel, O'Regan & Radach (2000) point out, although it "looks intelligent", it only does so because it provides the reader with, on average, the visual input when require (pp. 368, Deubel et al, 2000).

The forward progression of saccades can however be interrupted by foveal inhibition. Specifically, a saccade can be delayed by an inhibitory signal that originates from the lexical processing module, the enabling trigger for this inhibition relates to the lexical activity of the foveal word. This inhibitory mechanism is considered a central component of the SWIFT model and is reflected in the model's name "SWIFT", an abbreviation for "*S*accade generation *W*ith *I*nhibition of *F*oveal *T*argets". Despite delaying the onset of the next saccade, foveal inhibition is limited and cannot stall the next saccade indefinitely, quite the contrary. Richter et al (2006) report that despite the maximum foveal inhibition being almost linearly related to word length; the maximum foveal inhibition for the longest word in the Potsdam Corpus was only 65ms (based on the simulation by Kliegl, Grabner, Rolfs & Engbert, 2004).

With respect to saccadic programming, SWIFT incorporates a similar set of principles to those incorporated in the E-Z Reader model, which as discussed previously draws upon the research of Becker and Jürgens (1979). Like the E-Z Reader model, SWIFT incorporates two sequential stages of saccadic programming, with a labile followed by a nonlabile stage, the former of which can be cancelled, while the latter cannot. SWIFT conceptualises the labile stage as engagement of the oculomotor system; but with the final saccadic target not determined until this stage reaches completion. According to the architects of SWIFT, allowing a saccadic target to be determined at the end of the labile stage prevents the processing system from having to determine a saccadic target that will likely become redundant by the time it is eventually executed. Saccade execution takes, on average, 25ms in the model, during which time pre-processing is suspended due to saccadic suppression (Matin, 1974) and is only re-engaged 50ms after the eye re-stabilises to allow visual input to propagate to the visual cortex (i.e., the eye-to-brain lag; Foxe & Simpson, 2002; Reichle et al, 2003). It is assumed that lexical completion continues throughout a saccade's execution, as this stage operates independently from visual input.

The temporal evolution of a word's level of activation progresses in two discrete stages: the pre-processing and lexical completion stages. Given that increased levels of activation are associated with becoming a target for an upcoming saccade, SWIFT assumes that before a word is processed and once a word has been fully processed, its activation levels are set to zero. In-between these two extremes, pre-processing increases the activity level until it reaches its maximum - which is negatively related to a word's frequency - while the lexical completion stage allows activity to diminish. Finally, a global decay process that represents memory leakage acts to slowly reduce activation levels for all words at a constant rate.

Word frequency therefore imposes its influence in the model by modulating the maximum activation level a word can possess, with higher maximums related to increased difficulty and therefore attributed to lower frequency words. Since high levels of activation increase the probability of a word being chosen as a target for an upcoming saccade, this will result in low frequency words receiving more fixations and/or refixations than a high frequency word, thus affecting gaze duration for that word.

The model assumes a temporal asymmetry between these two stages of lexical processing, with the pre-processing stage ascending at a faster rate than lexical completion descends. Indeed, Richter et al (2006) report that pre-processing is approximately 90 times quicker than the lexical completion stage. This temporal asymmetry reflects the assumption that pre-processing is associated with the early stages of processing in which the word is included in the pool of potential targets for an upcoming saccade, and is therefore assumed to be a quicker process than the lexical completion stage, which reflects later memory retrieval processes. While the authors suggest that the pre-processing stage involves the extraction of preliminary information from the word, they do not preclude the possibility that semantic information can influence the efficacy of this stage.

A second variable that modulates both the pre-processing and lexical completion stages is a word's predictability. The model assumes that, providing a word is not being fixated, a highly predictable word will slow the accrual of activation during the pre-processing stage, thereby reducing the probability that it will enter the pool of potential targets. Like the E-Z Reader model, SWIFT therefore predicts that a word that is highly constrained by its preceding sentence context will be skipped more frequently than a word that is not, in agreement with previous research (e.g., Balota, Pollatsek & Rayner, 1985; Drieghe, Rayner, Binder, Ashby, & Pollatsek, 2001; Rayner & Well, 1996). Once a word reaches its lexical completion state, the processing rate for highly predictable words is increased, allowing them to be recognised faster, and therefore returning to a zero level of activation faster than an unpredictable word. Thus, SWIFT is also able to account for the finding that words that are low in predictability typically receive longer inspection times than words that are highly predictable (e.g., Balota et al, 1985; Rayner et al, 2001; Rayner & Well, 1996).

In SWIFT therefore, frequency and predictability act upon different processes. While a word's normative frequency influences a word's maximum level of activation, predictability influences the rate of processing. The architects of SWIFT suggest that this is a legitimate distinction given that predictability is independent of visual input (i.e., it is a top-down process that involves "guessing" the identity of upcoming words), while a word's normative frequency unfolds during the process of lexical identification (Engbert et al, 2005). Within the model, therefore, predictability has the potential to exert an earlier influence on the time course of lexical processing than a word's normative frequency.

Despite saccadic programming being independent from successful lexical identification (as with the E-Z Reader model), numerical simulations demonstrate that SWIFT is capable of reproducing word frequency effects in gaze duration (Engbert et al, 2005). This is possible because low frequency words will receive comparatively higher levels of activation compared to high frequency words, making them more attractive candidates for both fixations and refixations. Indeed, as will be recalled, low frequency words are skipped less frequently and receive more fixations than high frequency words (e.g., Henderson & Ferreira, 1993; McConkie, Kerr, Reddix, Zola & Jacobs, 1989; Rayner & Fischer, 1996; Rayner, Sereno & Raney, 1996). Engbert and Kliegl (2011) suggest that this process of saccade generation can explain the finding that difficult words do not merely increase single fixation durations on a word (which can occur when the foveal inhibition mechanism has been activated); rather, gaze durations are also increased via a series of multiple fixations.

It should be noted, however, that despite SWIFT's capability in simulating the frequency effect, the simulated effects are smaller in magnitude than the experimental data. Engbert et al (2005) acknowledge this discrepancy and attribute the difference to the fact that many other lexical variables that correlate with frequency (e.g., neighbourhood frequency) are not currently specified in the model.

SWIFT is also capable of predicting preview benefits, which as will be recalled, refer to the finding that fixation durations are longer on words deprived of a parafoveal preview compared to words that are not (e.g., Rayner, 1975; see above for more detail). Specially, all words within the effective span of apprehension can undergo lexical processing simultaneously. Therefore, if a word receives an invalid preview prior to fixation, the pre-processing stage for that word will have to wait until the word receives a direct fixation, increasing the chance that that word's activity levels will permit multiple fixations compared to a situation where a valid parafoveal preview had been displayed. Kliegl and Engbert (2003) demonstrated, using numerical simulations, that SWIFT (Version 1; Engbert et al, 2002) was capable of predicting preview benefit effects. By reducing a target's activation to zero upon a saccade into that target region, they found target word fixation durations increased, it was skipped less frequently and was more likely to attract a regressive fixation. Additionally, provided foveal inhibition is initiated quickly upon fixation, this mechanism could also act to lengthen a fixation on the target word following an invalid preview.

Spillover effects, in which foveal difficulty inflates fixation durations on subsequent words (e.g. Rayner & Duffy, 1986), can also be accommodated by SWIFT. Schad and Engbert (2012) note three sources responsible for their occurrence within the model. First, because the mechanisms involved in lexical processing are considerably slower than the brainstem circuitry responsible for saccade generation (Sparks, 2002), foveal inhibition will only take effect after some delay, which can often result in foveal inhibition only being realised after the next word in text receives a fixation. Second, due to acuity constraints, a long foveal word will slow parafoveal processing, and in so doing, increase the probability that once fixated, the parafoveal word will be refixated. Finally, difficult foveal words are likely to absorb the majority of the attentional resources available for processing words within the effective span of apprehension, thereby resulting in little pre-processing of parafoveal words compared to a situation where the foveal word was easy to process. The mechanism responsible for this modulation of attentional resource will be discussed next, within the related context of how foveal difficulty can modulate preview benefit (e.g., Henderson & Ferreira, 1990).

SWIFT III (Schad & Engbert, 2012) extends earlier versions by incorporating a ‘zoom lens’ mechanism (Eriksen & St. James, 1986) into the model. This mechanism allows attention, during the lexical completion stage of processing, to either dilate or contract in response to easy or difficult to process foveal words, respectively. This permits an account of Henderson and Ferreira’s (1990) finding that preview benefit can be modulated by foveal difficulty. Specifically, a difficult foveal word will cause attention to contract, and this will prevent attention spreading to the parafoveal word, minimising or preventing the accrual of parafoveal preview. If however, the foveal word is comparatively easy to process, the distribution of attention will be far less restricted, so parafoveal uptake can commence prior to direct fixation.

Engbert et al (2005) suggest that given the asymmetry of the eccentricity constraints imposed by SWIFT, the model naturally predicts that a reader should aim for a word’s preferred viewing location (Rayner, 1979), characterised by its leftward deviation from the centre of a word. As will be recalled, processing rate in SWIFT is conceptualised as an asymmetrical Gaussian distribution, which is calculated at the letter level. Since this distribution has a leftward shift, it makes sense that the oculomotor system should target the left of a words centre, as this position will be associated with the faster overall processing rate taking into account each of the letters within the word. This makes the preferred viewing location the most efficient place to fixate within a word. Therefore, in contrast to the E-Z Reader model, SWIFT does not conceptualise the asymmetric distributions with leftward tendencies as an effect solely caused by systematic undershoots of the target; rather, they suggest that it could simply be symptomatic of inherent eccentricity constraints.

SWIFT (version 2 onwards) also incorporates a saccadic error component. This component was an advance on the first version of SWIFT since it allowed the model to predict landing site distributions. Like the E-Z Reader model, SWIFT adopts McConkie et al’s (1988) assumptions regarding the inclusion of saccadic error inherent in the oculomotor system. In SWIFT (version 2 onwards), the executed saccadic length is the sum of the intended saccade length, plus systematic error, plus random error. The systematic error component incorporates under- and over-shoots of the target for saccades that travel longer or shorter than the preferred distance. The random error component is represented in the model as a Gaussian distribution with a mean of zero, the standard distribution of which increases with increasingly remote launch sites. By incorporating these principles, SWIFT is able to successfully simulate the obtained variability in landing site distributions (Engbert et al, 2005).

It follows from the preceding set of principles, therefore, that like the E-Z Reader model, SWIFT allows the possibility that some saccades will miss their intended target words (i.e. there will be ‘mislocated fixations’). However, unlike proposals associated with the E-Z Reader model, where such a scenario is predicted to result in a stay and process strategy, in the SWIFT model, these mislocated fixations initiate a quick error-correcting saccade to relocate the eye to a more optimal location. The position of the new saccade will not necessarily coincide with the original target word however, since the destination of a new saccade is only determined at the end of the labile stage of saccadic programming. At this stage, therefore, the original target may no longer be the word with the highest activation level, and as such, may no longer be the target for the 'corrective' saccade. Nevertheless, the initiation of the quick error-correcting saccade will act to reduce fixation durations near word boundaries, where mislocated fixations will be most prevalent[1].

Thus, SWIFT directly incorporates Nuthmann et al's (2005; 2007) conceptualisation of how the oculomotor system might respond to mislocated fixations, and in-so-doing, the model is able to simulate the approximate function of the IOVP effect as first reported by Vitu et al (2001). However, for SWIFT to model the magnitude of the IOVP effect found in experimental data, its architects (Engbert et al, 2005) incorporate a further assumption into their model that allows the duration of the nonlabile stage to be influenced by intended saccade length, a relationship that is not without precedence in the literature (e.g., Adams, Wood & Carpenter, 2000; Kalesnykas & Hallett, 1994;

[1] It is worth highlighting that the quick error-correction mechanism in SWIFT only applies to saccades that miss the intended target word and does not extend to within-word saccadic error.

Wyman & Steinman, 1973). Specifically, shorter saccades are associated with longer saccadic latencies. The addition of this moderator allows the SWIFT model to predict the peak in the IOVP distribution. As an aside, this added factor also provides the model with an explanation for parafoveal-on-foveal effects that stem from parafoveal word length (e.g., Drieghe, Brysbaert & Desmet, 2005), as saccades into long words will generally be longer and therefore associated with shorter saccadic latencies, reducing fixation durations on words preceding long parafoveal words.

Regarding parafoveal-on-foveal effects, it seems logical that a model advocating parallel lexical processing during reading should naturally be able to account for parafoveal-on-foveal effects. However, it will be recalled that foveal inhibition in SWIFT is only permitted to stall the progression of the eyes based on foveal, not parafoveal, difficulty. And in fact, SWIFT possesses no explicit mechanism capable of predicting parafoveal-on-foveal effects. While Engbert et al (2002) acknowledge that adjusting the model to allow inhibition by parafoveal targets would be possible, all versions of SWIFT retain the assumption that only foveal inhibition can delay the progression of the eyes.

Instead, Engbert and Kliegl (2011) suggest that parafoveal-on-foveal effects are predicted by the SWIFT architecture via the simultaneous activations of words, which in turn influences selection effects. Specifically, whether or not a saccade is executed from word n to word n+1 will depend upon both the activation level of word n+1 and the time spent viewing word n. If word n+1 is a low frequency word, it will be more likely to attract a fixation away from word n. While a high frequency parafoveal word will be less likely to attract a fixation, thereby increasing the probability of a refixation on word n. This variation in targeting decision is driven by low frequency words having a higher activation maxima compared to high frequency words. This pattern of events will become more likely the longer word n is fixated. A consequence of a parafoveal word n+1 causing a refixation is that the initial fixation in that case becomes re-labelled as the first (of multiple) fixations. All remaining initial fixations fall into the category of single fixation duration. Due to this partitioning of the data, single fixation durations will be shorter for high compared to low frequency words. While this explanation seems somewhat convoluted, Engbert and Kliegl (2011) note that it is based upon the simulations carried out by Engbert et al (2005) that showed parafoveal-on-foveal effects.

SWIFT can therefore be summarised using the following principles (1-7 taken directly from Engbert et al, 2005, pp781; 8 discussed in Schad & Engbert, 2012); each of which has been described and discussed above within the context of the model’s ability to account for the benchmark phenomena found in the eye movement record during reading:

1. Spatially distributed processing of an activation field
2. Separate pathways for saccade timing and saccade target selection
3. Random saccade generation with time-delayed foveal inhibition
4. Two-stage saccade programming with labile and nonlabile stages
5. Systematic and random errors influence saccade length
6. Error correction of mislocated fixations
7. Modulation of saccadic latency by saccade length
8. Zoom lens modulation of the attentional gradient.

1.4. Summary

The E-Z Reader and SWIFT models share many of the same principles. Both assume that word recognition progresses in two stages; both allow for frequency and predictability to influence the duration of word identification; both draw on Becker and Jürgens' (1979) research relating to saccadic programming; and both incorporate McConkie et al's (1988) principles of random and systematic saccadic error. There are therefore many parallels between the two models.

The major distinguishing feature, however, draws us back to the link between fixation location and attention during reading, with the cores of these models being built upon very different conceptualisations. While serial models, such as E-Z Reader, stipulate that attention can only be focussed upon one word at a time, processing gradient models, such as SWIFT, assume a much looser coupling between fixation location and attention, allowing multiple words to be lexically processed simultaneously.

As can been seen from the model descriptions above, both the E-Z Reader and SWIFT models are capable of providing explanations for a wide array of benchmark phenomena. It should be noted however, that in doing so, the models have had to evolve to a point where distinguishing the empirical consequences of their core assumptions has become a challenge. Consequently, to fully understand the nature of the relationship between fixation location and attention during reading, researchers have sought to devise various ways to investigate this long-standing question. A summary of these approaches and their relationship to this question will be considered in Chapter 2.

CHAPTER 2

Testing the Link between Fixation Location and Attention during Reading

Models of eye movement control have played a central role in driving the research into the link between fixation location and attention during reading. This research has typically focussed on effects for which serial and parallel models (such as E-Z Reader and SWIFT) make divergent predictions, such as semantic preview benefit, parafoveal-on-foveal effects and more recently, word n+2 preview benefit. This chapter will outline how the research in these areas has progressed, paying particular attention to how well the models – or at least their respective perspectives – can account for the research findings, and if not, how the models have needed to adapt in order to do so. Prior to outlining this research, however, the relatively uncontroversial aspects of preview benefit will be discussed in order to lay the foundation for what follows.

2.1. Preview Benefit

As discussed in Chapter 1, the boundary paradigm (Rayner, 1975) provides a gateway into investigations regarding how much pre-processing has occurred on a word prior to fixation. In this paradigm, the parafoveal preview of a target word is denied; only once the eye passes an invisible boundary located immediately before the target word does the preview switch to the target.

Participants are typically oblivious to these changes as they are triggered when the visual system is supressed during a saccade (Matin, 1974). Theoretically, therefore, this paradigm provides a measure of how much pre-processing is conducted on a target word before that word is directly fixated (see Chapter 1, Section 1.2.1 for a full description of the methodology). It should be noted, however, that while the processing ‘advantage’ of an identical over invalid preview is generally termed a preview ‘benefit’, recent research suggests that this temporal difference can be better understood in terms of both benefit *and* cost; the latter originating from interference related to the word change (Murray, Rayner & Wakeford, 2013; Risse & Kliegl, 2013; Schotter, Reichle & Rayner, 2014). This cautionary note aside, there is little doubt that preview benefit experiments have provided an invaluable tool for investigating the extent to which parafoveal processing occurs during ‘natural’ reading[2].

To investigate the levels of representation extracted from a word prior to fixation, the relationship between preview and target has been manipulated across a range of linguistic dimensions, including those associated with the early stages of word recognition (such as the extraction of orthographic and phonological codes) as well as those associated with the later stages of word recognition (e.g., the semantic code); these will now be considered in turn.

[2] To be exact, gaze contingent display change experiments such as these can only be said to approximate natural reading since the word change might, albeit at a subconscious level, influence attentional processes.

2.1.1 Orthographic and Phonological Preview Benefit

There is an abundance of evidence for the existence of orthographic preview benefit, in which fixation durations are higher following previews that do not share the same letter identities as the target word, compared to previews that do (e.g., Balota, Pollatsek & Rayner, 1985; Inhoff, 1989; Rayner, 1975; Starr & Inhoff, 2004; White, Rayner & Liversedge, 2005). McConkie and Zola (1979) further demonstrated that the advantage of receiving a correct preview does not seem to be driven by the integration of perceptual information across fixations. They asked participants to read sentences that were displayed in AlTeRnAtInG cAsE, these presentations either remained static between fixations (e.g., CaSe -> CaSe) or switched (e.g., cAsE -> CaSe). McConkie and Zola reported no evidence that case switching had a disruptive effect on reading performance compared with when the display remained static. Indeed, they reported that not one participant (out of eight) noticed the display change occurring. It appears therefore that while featural information is extracted prior to fixation, this information is quickly converted into abstract letter codes, and it is these that are retained from fixation to fixation.

There is also evidence for the existence of a phonological preview benefit, in which target word inspection time is shorter if the preview contains a phonologically related, compared to phonologically unrelated preview (e.g., Ashby & Rayner, 2004; Ashby, Treiman, Kessler & Rayner, 2006; Miellet &

Sparrow, 2004; Pollatsek, Lesch, Morris & Rayner, 1992). These effects are apparent whether previews consist of homophones (e.g., sent -> cent; Pollatsek et al, 1992) or pseudohomophones (e.g., roze -> rose; Milliet & Sparrow; 2004). Ashby et al also found a target duration advantage for previews that contained vowel concordant phonemes (e.g. chirp -> cherg) compared to vowel discordant phonemes (e.g., chirp -> chorg). Since effects of phonological preview benefit have been reported in French (Milliet & Sparrow, 2004), it is apparent that these effects are not restricted to the English writing system. It seems therefore that it is the letter string phonology that is activated prior to fixation rather than a word's lexical entry *per se*.

Effects of orthographic and phonological preview benefit pose little difficulty for either the SWIFT or to the E-Z Reader models of eye movement control. It will be recalled from Chapter 1 that according to SWIFT, all words falling within the effective span of apprehension can undergo pre-processing prior to fixation; the model therefore predicts orthographic and phonological preview benefits. It will also be recalled that the E-Z Reader model posits two stages of lexical processing: L1 and L2; the completion of L1 triggers a saccade to be programmed to the next word in text, while the completion of L2 triggers an attention shift to the next word. Therefore, providing the time required to complete the second stage of lexical processing on the fixated word completes faster than the time required to execute a saccade to the parafoveal word, a shift of attention to that word should occur and parafoveal preview benefit should be obtained. Given the tight time constraints on parafoveal processing that this architecture imposes, however, it is generally considered that only the early stages of word recognition should occur with a parafoveal word prior to its fixation. Both the orthographic and phonological codes are considered to be examples of such early processes (at least in alphabetic scripts: Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001).

2.1.2. Semantic Preview Benefit

2.1.2.1. Semantic Preview Benefit: The Models

In contrast to the extraction of orthographic and phonological codes, the extraction of a semantic code is generally considered to reflect a late stage of lexical access and therefore typically requires more time to become activated (e.g., Coltheart et al, 2001). Given the architectures of the SWIFT and E-Z Reader models outlined in the preceding Chapter, it is perhaps not surprising that evidence pertaining to semantic preview benefit has become a hotly debated topic.

To be clear, for the reasons outlined above, Processing Gradient models can account for semantic preview benefit, at least at a theoretical level. However, for the effects to arise within the E-Z Reader architecture, enough cases must arise where semantic activation of the parafoveal word completes (a) *after* the completion of L2 on the foveal word, but *before* a saccade out of it is executed and (b) not so quickly that the parafoveal word is skipped (i.e., before the completion of M1 on the foveal word). As Radach and Kennedy (2013) note, the situations in which these conditions are satisfied should be extremely rare; therefore, these effects should not typically arise within a serial architecture[3]. Given that Processing Gradient models can, in principle, account for semantic preview benefit, while serial models such as E-Z Reader typically do not, this presents an interesting platform on which the two models of eye movement control can be tested.

2.1.2.2 Semantic Preview Benefit: The Evidence

Studies investigating semantic preview benefit have typically manipulated the sematic relatedness of the preview and the target word, on the basis that responses to semantically related word pairs are facilitated compared to unrelated pairs (Meyer & Schvaneveldt, 1971; also see Neely, 1991). By extension, it is suggested that semantically related previews should facilitate target word identification compared to unrelated previews. Using the boundary paradigm, Rayner Balota and Pollatsek (1986) tested this hypothesis and asked participants to read sentences such as *"My younger brother has brilliantly composed a new song for the school play"*, in which the pre-fixation preview of "song" was either "song" (valid), "tune" (related), "door" (unrelated), or "sorp" (a visually similar nonword). Only once the eye passed an invisible boundary, located before the critical word, did the target word

[3] However, Schotter, Reichle and Rayner (2014) have recently proposed that the E-Z Reader can in fact account for semantic preview benefit; the relative merits of this modelling exercise and the assumptions that underlie it will be returned to in Chapter 5.

“song” appear. Despite showing that their critical words produced facilitation in a classic priming experiment, Rayner et al found no evidence for a semantic preview benefit during reading. Confidence in this null result was recently reinforced by Rayner, Schotter and Drieghe (2014) who again observed no evidence for a semantic preview benefit after running an almost direct replication of Rayner et al’s experiment.

However, in the example sentence provided by Rayner et al, it can be seen that the word to the left of the target contains only three letters, and as short words are frequently skipped (Rayner & McConkie, 1976), the prior fixation may in fact have fallen two words to the left of the target, seriously reducing the chance of it eliciting a semantic preview benefit. A more general problem with experiments investigating semantic preview benefit using associative previews is that while there may be semantic facilitation, there is also a word change that might be expected to give rise to some form of inhibition. Semantically related word pairs, such as north–south, rattle-bottle and arms–legs, (from Rayner et al, 1986), all have very different meanings, and this could result in an inhibitory effect on on-going sentence interpretation. Rayner et al (1986) attempted to test this possibility by asking participants to rate their sentence pairs for similarity of meaning and reanalysing the results from only the 20 sentence pairs rated as most similar in meaning. Since this analysis again failed to show a semantic preview benefit, they dismissed this as an explanation for their null result. However, a measure of overall sentence meaning does not necessarily capture the extent to which a local change in word meaning might have disrupted the reading process at the point at which it was first encountered. Therefore, the suggestion that the word change may have given rise to some form of interference must remain a possibility.

Addressing this issue, Altarriba, Kambe, Pollatsek and Rayner (2001) also used the boundary paradigm but with fluent English-Spanish bilinguals. They employed semantically-related previews which were translations with virtually the same meaning as their targets, thereby reducing the possibility of interference. All changes involved a word preview from the other language that was either: *cognate* (orthographically and semantically similar), *noncognate* (semantically similar but orthographically dissimilar), *pseudocognate* (semantically unrelated but orthographically similar), or a *control* (unrelated orthographically and semantically) to the target word. They found no evidence for a semantic preview benefit in the absence of orthographic similarity. However, it remains possible that facilitation might not cross over between the lexica of the two languages, and as Hohenstein, Laubrock and Kliegl (2010) point out, since the previews and targets were in different languages, switching costs (Meuter & Allport, 1999) might mask any semantic preview benefit that might have been accrued.

Hyönä and Häikiö (2005) hypothesised that the lack of evidence from previous experiments for a semantic preview benefit could be due to the semantic manipulations not being sufficiently strong to elicit an effect. They drew on the research of Calvo and Castillo (2005), who found evidence (from a non-reading task) to suggest that threat-related parafoveal words have “privileged access” compared to parafoveal neutral or positive words. Like these other studies, Hyönä and Häikiö employed the boundary paradigm and presented participants with sentences like: “*In my opinion, any animal’s cub is extremely cute”,* in which the target (cub) received either an identical preview (e.g., cub) a negatively valenced preview often involving obscene or curse words (e.g., penis) or a neutral preview (e.g., penny); sentences were presented in Finnish, this example represents a translation. Under these conditions, there was no evidence - in either durational measures or pupil dilation recordings - that the negatively valenced previews were semantically processed prior to fixation, with no discernable differences between them and the neutral previews when the target was eventually fixated. Hyönä and Häikiö suggest that their failure to replicate Calvo and Castillo’s result of privileged access for threatening words in the parafovea could be related to task differences. Indeed, there are many extra processing demands associated with reading that are not present in other non-reading tasks such as lexical decision. Therefore, despite Hyönä and Häikiö avoiding the confound that a semantically related word may cause interference (there was no semantically related condition), they obtained no evidence to suggest that the semantic code of a parafoveal word was extracted prior to fixation.

Hohenstein et al (2010) reasoned that these early failed attempts to uncover a semantic preview benefit might have been driven by a lack of control over the target word's preview duration. In boundary paradigm experiments such as these, the temporal availability of a preview will always be a function of the preceding fixation/s duration and therefore inherently uncontrollable. Hohenstein et al's reasoning was borne out of research conducted by Sereno and Rayner (1992) who used fast priming to demonstrate that semantic priming of a *foveal* word is sensitive to prime duration. Sereno and Rayner presented participants with sentences such as "*He cleaned his pipes once a month"* where the target word (e.g., pipes) received an invalid preview (e.g., gzsd) prior to passing an invisible boundary located before the penultimate letter of the pre-target word. Upon passing the invisible boundary, either a semantically related (e.g., cigar), or unrelated (e.g., witch) prime was briefly displayed before the target word (e.g., pipes) was eventually presented. Sereno and Rayner obtained shorter target word inspection times following a semantic prime at a duration of 30ms, but this effect disappeared for durations of 21ms and 39ms. This sensitivity to target prime duration implies that the semantic code takes some minimal time to accrue, but should too much time lapse, the derived code will begin to interfere with target word processing.

To investigate the possibility that parafoveal prime duration may be responsible for the lack of semantic preview benefit, Hohenstein et al employed the fast priming technique, but instead of controlling foveal prime duration, they controlled *parafoveal* prime duration. Conducted in German, participants were exposed to sentences like "*With the excavation, skulls had shown up"* in which the target "skulls" received either a semantically related (e.g., bones) or unrelated (e.g., boots) prime. Upon fixation of the *pre-target* word, the target word preview changed from a nonword to either a semantically related or an unrelated prime for durations of either 35ms, 80ms or 125ms; after which point the target word was displayed. Hohenstein et al report a significant semantic preview benefit when the prime duration was 125ms, with shorter gaze duration following the related compared to the unrelated primes (Experiments 1 & 2; effect sizes in gaze duration of 23ms and 21ms, respectively). When primes were presented in bold typeface, an 18ms effect, significant in gaze duration was observed, but only with prime durations of 80ms (Experiment 3). The authors interpret these results as representing the minimal time a parafoveal prime must be present for the semantic code to become activated, a process that can become more efficient with primes of increased saliency. Further, they suggest that when the prime is visible for too long in the parafovea (under the increased saliency condition at 125ms), the prime can act to disrupt target word processing. While this experiment would benefit from replication, especially given the transient nature of the semantic preview benefit between experiments, it does fit nicely with the premise that semantic associates can interfere with target word processing.

Perhaps the most convincing evidence to date that a semantically related preview can (a) interfere with target word processing when there is discord between preview and target word meanings, and (b) facilitate target word processing when the two words share a common meaning, comes from Schotter (2013). Schotter proposed that meaning changes between semantically related previews and targets might potentially eliminate any evidence of semantic preview benefit. To test her theory, she employed the boundary paradigm but differentiated between synonymous previews (e.g., movie-video) and semantic associate previews (e.g., north-south). Participants were exposed to sentences like "*My friends have the same favourite movie that they watch every week*", in which the target word preview (movie) was either a synonym of the target (e.g., video), an unrelated preview (e.g., water), or a semantic associate (e.g., audio; Experiment 2 only). Gaze duration on the target following a synonymous preview was significantly shorter than when it followed an unrelated preview (Experiments 1 & 2; effects sizes 16ms and 9ms, respectively). Importantly, there was no evidence for a semantic preview benefit when the semantically related and unrelated preview conditions were compared (Experiment 2), replicating earlier studies failure to obtain semantic preview benefits with semantic associates (e.g., Rayner et al. 1986).

Taken together, these results demonstrate that while a semantic code can be extracted from a parafoveal word prior to fixation, this only occurs when the overlap in meaning between the preview and target is sufficiently high. Indeed, supplementary analyses revealed that the degree to which sentences were rated as similar in meaning (in a prior norming exercise) was negatively correlated with time spent viewing the target word, suggesting that the changes in word meaning that result from using semantic associates is likely to be responsible for the failure of previous efforts to obtain a semantic preview benefit.

There is, however, reason to doubt that a close correspondence between preview and target word meaning is a necessary prerequisite for semantic preview benefit to be expressed: semantic preview benefits originating from semantic associates, not synonyms, have been observed under optimal conditions. The first boundary experiment to obtain a semantic preview benefit using semantic associates involved 2-constiuent Finish compound nouns, in which the compound contained pre-target (first constituent) and target (second constituent) regions. Conducted by White, Bertram and Hyönä, 2008, this design was inspired by the finding that larger preview benefits are typically observed when a space does not separate the pre-target and target regions (e.g., Hyönä, Bertram & Pollatsek, 2004). By extension, therefore, semantic pre-processing may be more likely under these conditions too. White et al presented participants with 2-part Finnish compound nouns embedded in Finnish sentences like *“According to Laura vanillasauce goes together well with apple pie”* in which the preview of the second constituent was either identical to the target (e.g., vanillasauce), semantically related (e.g., vanillamustard), semantically unrelated (vanillapriest), or a pronounceable nonword preview (no example provided). The boundary was located between the penultimate and last character of the first constituent. Their results showed no statistical evidence for a semantic preview benefit when either of the two constituents were analysed separately, however, when analysed as a combined region[4], inspection time following a semantically related preview was 33ms faster than when following a semantically unrelated preview. While the precise explanation for why this effect was only apparent in a relatively late measure is unknown (perhaps the semantic activation was buffered until required for whole-word interpretation), this study provided an important first step in establishing that a semantic code can be extracted prior to fixation, albeit under these optimal proximal conditions.

Another contradiction to the proposal that interference is responsible for the elusive nature of semantic preview benefit was recently published by Hohenstein and Kliegl (2013), who report obtaining semantic preview benefit using semantic associates in German. They conducted a series of experiments with materials very similar to their previous study discussed above, presenting participants with sentences like "*With the evacuation, bones came to light*", in which the target "bones" received either a related (e.g., skulls) or unrelated (e.g., boots) prime. This time, however, rather than controlling the duration of

[4] Using a measure that included all fixations on the compound after the second constituent was first entered and before it was exited to the right.

parafoveal preview with the fast-priming technique, they used the standard boundary paradigm (Rayner, 1975) with the previews visible until the eye passed an invisible boundary located before the target word. Under these conditions, Hohenstein and Kliegl report obtaining a semantic preview benefit in three separate experiments, with their respective effect sizes being 31ms, 18ms, 27ms; the latter two values are collapsed over whether or not the first letter of the noun was capitalised (Experiment 2) and whether the contrast between text and background was set to normal or low (Experiment 3). The semantic preview benefit was obtained regardless of noun capitalisation or contrast setting. In addition, their extensive analysis revealed evidence to suggest that the semantic preview benefit was moderated by (a) launch site distance from the target and (b) pre-target gaze duration.

Hohenstein and Kliegl's (2013) results are interesting since they contrast not only with their own previous efforts to obtain semantic preview benefit, where the effects were only present during certain prime durations (Hohenstein et al, 2010), but also the results of Schotter (2013), which suggest that semantic preview benefits should not occur using semantic associates. Hohenstein and Kliegl attribute the divergence between their 2010 and 2013 results as reflecting the "different mechanisms" that are at work in the parafoveal fast-priming and boundary experiments (Hohenstein and Kliegl, 2013; pp173). Indeed, Hohenstein et al's (2010) preview-to-target switch occurred while the eye was stationary, which may have caused attention to be distributed in an atypical manner in those trials.

Schotter (2013) proposed her own explanation for why Hohenstein and Kliegl (2013) obtained semantic preview benefit with semantic associates, while her research suggests that this should not typically occur. Specifically, Schotter does not allude to interference as a cause for prior failures in obtaining preview benefit; rather, she suggests that the reason may be routed in the high processing demands inherent in reading English. Schotter suggests that the German language has more efficient connections between orthography and semantics, while English typically requires phonological decoding prior to the semantic codes becoming activated, and it is these extra processing demands in English that prevent semantic preview benefit from occurring. Indeed, Hohenstein agrees that the mediation of phonological encoding creates language dependent differences (e.g., Laubrock & Hohenstein, 2012). The reason Schotter suggests she was able to obtain a semantic preview benefit using synonyms is that synonyms may either be “stored together”, or share “stronger connections” than words that are simply semantic associates (Schotter, 2013, pp629). Therefore, the differences between Schotter’s results and those originating from the Hohenstein laboratories may be due to language related differences.

Hohenstein and Kliegl (2013) suggest another (although arguably related) explanation for the differences: their use of a high frequency pre-target word may have allowed attentional resources to be distributed more freely in their experiment (e.g., Henderson & Ferreira, 1990; White et al, 2005). A further explanation could simply be that Hohenstein and Kliegl created preview-target pairs that were more closely related than in previous studies. In their example above, “bones” and “skulls” carry very similar meanings within the context of the sentence and it is perhaps this similarity of contextualised meaning that drives their ability to obtain a significant preview benefit: the nature of the change from the preview here seems less prone to possible meaning inference than in previous attempts to uncover a semantic preview benefit (e.g., Rayner et al, 1986).

2.1.2.3 Semantic Preview Benefit: Summary

To conclude, evidence pertaining to a semantic preview benefit is becoming increasingly difficult to refute, with several studies now reporting such effects (Hohenstein et al, 2010; Hohenstein & Kliegl, 2013; Schotter et al, 2013; White et al, 2008). Schotter suggests that her results can be reconciled with the E-Z Reader model, owing to the privileged connections that she suggests exist between synonymous words, which result in a temporal efficiency that falls within the time constraints outlined by the E-Z Reader model. Whether Schotter’s suggestion is valid, or whether the differences between experimental findings are related to item selection effects, language related differences, or whether interference might play a role are all possibilities that are, at present, unresolved. Gaining a better understanding of the role that interference might play in previous failed attempts to uncover semantic preview benefit may provide some insight into these unanswered questions. The role that interference might play in expression of sematic preview benefits will be returned to in Chapter 5.

2.1.3 Preview Benefit Summary

It should be clear from the above discussion that while orthographic and phonological preview benefit are uncontroversial and easily accounted for by both serial and parallel models of eye movement control, the debate on semantic preview benefit remains far from settled. While effects of semantic preview benefit have been observed, it remains an open question as to whether previous divergent findings might be related to differing levels of interference between previews and targets, or whether other explanations might be implicated.

One final question related to the occurrence of semantic preview effects is whether they might be accounted for by saccadic undershoots followed by a stay and process response (e.g., Drieghe et al, 2008; Rayner et al, 2004). According to this proposal, the eye may have targeted word n, but due to a saccadic undershoot, fall short, erroneously fixating (or refixating) word n.

According to the stay and process proposal, rather than relocating the fixation to the intended target location, word n+1 (with its preview still in place) is processed from this erroneous location. This explanation for semantic preview benefit is thus compatible with the E-Z Reader model since word n+1 is not processed before word n has been lexically identified. The plausibility of this proposal in accounting for semantic preview benefit will be explored directly in Chapter 5, while the general principle that a mislocated fixation is likely to be followed by a stay and process response will be returned to in Chapter 6.

2.2 *Parafoveal-on-Foveal Effects*

As will be recalled from Chapter 1, parafoveal-on-foveal effects refer to instances in which the properties of a not-yet-fixated word to the right of fixation influence foveal inspection times[5]. Interest in parafoveal-on-foveal effects stemmed not least from the perceived implication their existence would have for models of eye movement control, with positive evidence suggesting that information from multiple words was being extracted in a parallel fashion. However, as reports of parafoveal-on-foveal effects began to filter into the literature, the E-Z Reader model was able to evolve, and in doing so, highlighted how a serial perspective could be maintained in spite of these apparent parallel effects. This section will present the research that

[5] Since a word to the right of fixation may fall within foveal view (i.e., the central region in which visual acuity is sharpest), the label "parafoveal-on-foveal effect" is therefore one of convenience and should not be taken in its literal sense.

encouraged this evolution and discuss whether the auxiliary assumptions that E-Z Reader has incorporated provide a plausible explanation for the evidence.

Broadly speaking, there are two main classes of parafoveal-on-foveal effect: those that stem from the orthographic properties of a word to the right of fixation, and those that are lexical or higher level in nature. Prior to discussing the evidence for these effects, I will briefly mention one other effect that Drieghe (2011) likens to a parafoveal-on-foveal effect; namely, the influence parafoveal word skipping has on foveal inspection times.

2.2.1. Fixations Prior to Word Skipping

Several studies have shown that the decision to skip a parafoveal word influences foveal inspection times; the expression of these effects has however been inconsistent. While some researchers report increased fixation durations prior to word skipping (e.g., Drieghe, Rayner & Pollatsek, 2005; Hogaboam, 1983; Pynte, Kennedy & Ducrot, 2004; Rayner, Ashby, Pollatsek & Reichle, 2004), others do not (e.g., Drieghe, Brysbaert, Desmet, & Debaecke 2004; Radach & Heller, 2000). Research carried out by Kliegl and Engbert (2005) suggests that such inconsistencies might be driven by variation in item sets. Employing the Potsdam Corpus, in which 222 people read 144 German sentences, Kliegl and Engbert analysed foveal inspection times as a function of parafoveal word skipping while controlling for (a) localised text difficulty, (b) within-word fixation location prior to skipping, and (c) individual differences in reading rate. Results from this large corpus of eye movement data showed that short and high frequency words received shorter pre-skip fixation durations (i.e., pre-skip benefit) while long and low frequency words received longer pre-skip fixation durations (i.e., pre-skip cost) compared to a baseline of no skipping.

The most recent version of the SWIFT model (Schad & Engbert, 2012) is capable of replicating Kliegl and Engbert's (2005) findings. It is apparent that the model will predict skipping costs, given longer fixation durations on a foveal word - which could occur due to foveal inhibition or fluctuations in the stochastic random timer - since these allow more parafoveal processing to occur and thereby increase the probability that the parafoveal word will be skipped. In their publication of SWIFT-3, Schad and Engbert (2012) also report that the SWIFT model replicates Kliegl and Engbert's finding of skipping benefits. Schad and Engbert suggest that the process by which it did this was clearly linked to the zoom lens mechanism, as disabling it in a simulation on shuffled text caused the pattern of effects to revert back to skipping costs alone. The authors commit to focussing on the link between the zoom lens mechanism and skipping costs and benefits in a future publication.

In the E-Z Reader model, modulation of pre-skip fixation duration results from the assumption that the next word in text will always be the target of a saccade. So, for a word skip to occur, a saccade to the next word needs to be cancelled and a new one initiated, the process of doing so incurs a time penalty, explaining longer fixations prior to word skipping. While the E-Z Reader model is capable of predicting skipping costs, it is not clear how it can account for the skipping benefits reported by Kliegl and Engbert (2005). Further research on this topic is clearly required, and indeed, if Kliegl and Engbert's results can be replicated, then they will undoubtedly present a challenge to the E-Z Reader model.

2.2.2. Low Level Parafoveal-on-Foveal Effects

2.2.2.1. Low Level Parafoveal-on-foveal Effects: The Evidence

Orthographic parafoveal-on-foveal effects refer to cases in which the orthographic properties of a parafoveal word influence fixation duration on the currently fixated word. Typically, these effects have been investigated using the boundary paradigm (Rayner, 1975), in which word n+1 receives an invalid, typically orthographically irregular, preview prior to fixation, but with the fixation durations on word n being the measure of interest.

The first study to uncover such effects was conducted by Rayner (1975). Using the boundary paradigm, he presented participants with critical words that contained either an alternative word preview or a nonword preview, the latter typically involved orthographic illegality[6]. Rayner reported a significant

[6] Rayner employed three different types of nonword previews; however, these previews did not differ significantly from one another in the results with respect to parafoveal-on-

increase in last fixation duration when a nonword preview was presented to the right of fixation, but only when that last fixation fell within three character spaces of the critical word. The existence of orthographic parafoveal-on-foveal effects was thus demonstrated in the very first boundary paradigm experiment *providing* the eyes fell close enough to the parafoveal word.

Inhoff, Starr and Shindler (2000) also uncovered an orthographic parafoveal-on-foveal effect using the boundary paradigm; furthermore, this effect was not restricted to instances in which fixations fell close to the critical word. They found a 34ms increase in gaze duration when the word to the right of fixation was an illegal nonword preview (e.g., "qvtqp"), compared to an identical preview (e.g., "light"). They also obtained evidence that parafoveal visual distinctiveness influenced foveal inspection times, with gaze duration increasing by 25ms when the parafoveal preview contained a capitalised letters (e.g., "LIGHTS") rather than an identical (i.e., lower case) preview[7]. These results suggest that both orthographic irregularity and visual distinctiveness of parafoveal words can influence foveal inspection time.

The results of several subsequent studies have increased confidence in the validity of orthographic parafoveal-on-foveal effects (e.g., Angele et al, 2011; Starr and Inhoff, 2004), although, as observed by Rayner (1975), these

---

foveal effects. In an example set of items, 12 out of 15 of the nonword previews were orthographically illegal.

[7] Both of Inhoff et al's results were significant across a range of word-based measures on the pre-target word and were not restricted to gaze duration.

effects occasionally only manifest when fixations on the foveal word fall close to the parafoveal word (e.g., Drieghe, Rayner & Pollatsek, 2008). It should be stated, however, that a failure to obtain such effects remains prevalent in the literature (e.g., Angele et al, 2008; Rayner et al, 2007; White and Liversedge, 2004, White et al, 2005). Indeed, null effects have even been reported when the manipulation occurs on the second part of a two part compound word, where visual acuity should have been optimal and parallel processing encouraged (e.g., Hyönä, Bertram and Pollatsek, 2004).

Despite this inconsistency, there is a general consensus that these effects are real (e.g., Radach and Kennedy, 2013; Schotter et al 2012), cemented by the findings of Kliegl et al (2007) who – as will be discussed in the n+2 preview benefit section of this chapter – obtained evidence for long range orthographic parafoveal-on-foveal effects stemming from an illegal preview of word n+2 (i.e., a word two words to the right of fixation).

There is also evidence to suggest that sub-lexical properties of parafoveal words can influence foveal inspection times (e.g., Pynte, Kennedy and Ducrot, 2004; Underwood, Binns & Walker, 2005; White, 2008). Pynte et al (2004) report three gaze contingent experiments in which word n+2 was manipulated prior to fixation such that it contained either an identical preview, or one that differed by a single character to create a typographical error; with 'errors' always occurring near the beginning of word n+2. They found that when the error resulted in orthographic illegality there was a 13% reduced probability of skipping word n+1 (either “de” or “du”), suggesting a more cautious reading style had been adopted; word n inspection times were however unaffected (Exp1). When the typographical error created a legal nonword, there was no effect of word n+1 skipping, suggesting the previous effect had been triggered by the preview’s orthographic illegality. Gaze durations on word n were however 39ms shorter when the preview contained an error (Exp2). Finally this latter effect appears to have been driven by a lack of control over the typographical error previews and their targets. When trigram frequency was controlled (Exp3), there was no effect of the lexical status of word n+2 on fixations on word n. However, word n first fixation durations were 24ms shorter when the preview of word n+2 contained a constrained initial trigram - classified as ‘unfamiliar’ - compared to when the previews were unconstrained and therefore ‘familiar’[8]. Thus, the presence of unfamiliar trigrams in the parafovea appears to have attracted attention to the region of difficulty, resulting in reduced foveal inspection times. Kennedy (1998) and Hyönä and Bertram (2004) have both proposed that an ‘attraction’ mechanism might be responsible for such a process.

Taken together, these effects appear to be either orthographic (Exp1) or sub-lexical (Exps 2 & 3) in nature. Pynte et al’s finding of a reduced tendency

[8] Constrained trigrams are considered as ‘unfamiliar’ since only 1 other word shared the initial trigram, while unconstrained trigrams shared the same initial trigram with 25 other words, increasing their familiarity.

to skip word n+1 when word n+2 was orthographically illegal is similar to the skipping pattern reported by Radach, Glover and Vorstius (2007), who also reported a reduced probability of skipping word n+1 in the presence of word n+2 illegality. However, both these patterns are qualitatively different to the pattern observed by Angele and Rayner (2011) who report an *increased* probability of skipping word n+1 in the presence of word n+2 illegality; these studies will be discussed fully in Section 2.3.2 (Word n+2 Preview Effects: The Evidence). For now, however, it is important to note that divergent results such as these are not an uncommon feature of parafoveal-on-foveal effects between - and sometimes even within - experiments. Various explanations have been considered for such apparent inconsistencies. For example, evidence has converged from both reading and other types of task to suggest that foveal and parafoveal word length may play a role in the expression of parafoveal-on-foveal effects (e.g., Hyönä & Bertram, 2004 and Kennedy, Pynte & Ducrot, 2002). In addition Kennedy and Pynte (2005) report language dependent variation in the expression of parafoveal-on-foveal effects with differences between French and English related to the divergent results. Divergent results could also be related to item selection effects, among a host of other possibilities. Indeed, the above three studies differed from one another on all of these dimensions.

Inconsistent expressions of effects is also characteristic of sub-lexical parafoveal-on-foveal effects when durational measures are considered. For example, sub-lexical parafoveal-on-foveal effects do not always manifest in shorter foveal fixation durations. When such effects are present, they occasionally reflect an *increase* in foveal inspection times. For example, White (2008) obtained longer foveal fixation durations when a parafoveal word was orthographically unfamiliar compared to familiar. Underwood et al (2000) also reported increased foveal fixation durations when the parafoveal word contained an unfamiliar rather than familiar initial-trigram. Despite these inconsistencies, the fact that orthographic and sub-lexical parafoveal-on-foveal effects exist at all has proved to be a key factor in driving the development of models of eye movement control.

2.2.2.2. Low Level Parafoveal-on-foveal Effects: The Models

The SWIFT Model

Since the activation of the inhibition mechanism in the SWIFT model is restricted to foveal input alone, this model contains no dedicated mechanism to account for parafoveal-on-foveal effects. But since the saccadic targeting mechanism is influenced by the simultaneous activations of multiple words that dynamically evolve over time, simulations show that SWIFT can account for parafoveal-on-foveal effects. Specifically, a decision regarding whether to saccade out of a word - which necessarily influences foveal inspection time - will depend on a combination of (a) time spent inspecting the foveal word, and (b) the relative activation of all words falling within the perceptual span, including the parafoveal word (Engbert et al, 2005; Engbert & Kliegl, 2011, Schad & Engbert, 2012). Risse, Hohenstein, Kliegl and Engbert (2014) also report that SWIFT 3 is capable of simulating a numerical trend towards a word n+2 parafoveal-on-foveal effect. Therefore, as would be expected from a parallel model of eye movement control, SWIFT 3 does an appropriate job of predicting orthographic parafoveal-on-foveal effects. Whether or not the SWIFT model is capable of simulating the diversity of parafoveal-on-foveal effects discussed in the preceding section, does however remain to be established.

The E-Z Reader Model

Early demonstrations of orthographic and sub-lexical parafoveal-on-foveal effects were considered problematic for the E-Z Reader model. However, the introduction of the low level attentional scan in E-Z Reader 7 (Reichle et al, 2003) permits these effects to occur without violating that model's strict serial sequential assumptions. While this mechanism was originally designed to register the spatial coordinates for upcoming saccades, because it extracts information in parallel (unlike the serial operations of lexical processing), it provides a possible means through which low level orthographic irregularity in the parafovea might be detected prior to fixation, thereby influencing saccadic latencies (Rayner et al, 2003; Reichle et al, 2003). What is perhaps problematic about this explanation is that it provides little in the way of explanation for why parafoveal irregularities occasionally induce longer fixation durations on foveal words while other studies report a shortening in foveal inspection times. Nor does it explain the divergence in the word n+1 skipping rates reported by those manipulating the properties of word n+2 (e.g., Angele et al, 2011; Pynte et al, 2004). Nonetheless, in light of the general consensus that these effects are no longer able to differentiate between the processing gradient and serial classes of models, discussion will be re-directed to lexical parafoveal-on-foveal effects, as these effects cannot be explained by a low level attentional scan.

2.2.3. Higher Level Parafoveal-on-Foveal Effects

2.2.3.1. Lexical Parafoveal-on-Foveal Effects: The Evidence

Kennedy (1998; 2000) and Kennedy et al (2002) were the first to report evidence pertaining to lexical parafoveal-on-foveal effects. In the 2002 study, which was conducted in French, participants were required to read sequences of five words carefully but normally and to indicate upon completion of each sequence whether or not an article of clothing was present. Foveal (word 3) and parafoveal (word 4) words were either long (9 letters) or short (5 letters) and either high or low frequency. Initial letter trigram of the parafoveal word was also manipulated such that it was either informative or uninformative[9]. Kennedy et al report a complex pattern of results, but critically for the present

[9] Informative beginnings occur in relatively few lexical items, whereas uninformative beginnings are more common.

discussion, they report obtaining unorthodox frequency-driven parafoveal-on-foveal effects[10] when the foveal word was short and the parafoveal word was informative; the effect apparently driven, in part, by refixations. This effect was not present when the foveal word was long, with only the sub-lexical property of informativeness achieving significance.

A number of researchers (e.g., Pollatsek et al, 2006; Rayner et al, 2003; Schotter et al, 2012) have, however, argued that early reports of frequency-driven parafoveal-on-foveal effects, such as by Kennedy et al, may not translate to more 'natural' reading tasks. Indeed, when participants are simply required to read sentences (and answer the occasional comprehension question), frequency-driven parafoveal-on-foveal effects have proved rather elusive (e.g., Angele et al 2008; Carpenter & Just, 1983; Henderson & Ferreira, 1993; Inhoff, Starr, et al 2000). Furthermore, when these effects do arise in more natural reading tasks, their expression is often inconsistent (e.g., Hyönä & Bertram, 2004).

Hyönä and Bertram (2004) provide perhaps the most comprehensive set of experimentally obtained results related to parafoveal-on-foveal effects within a natural reading environment. Conducted in Finnish, their parafoveal word (n+1) was always an unspaced compound, in which either the initial

[10] An orthodox parafoveal-on-foveal effect refers to effects that mirror those obtained on the word undergoing the manipulation. For example, if the parafoveal word is a low frequency word, an orthodox parafoveal-on-foveal effect will manifest itself as increased foveal inspection times; the opposite being true for unorthodox parafoveal-on-foveal effects, where a low frequency parafoveal word reduces foveal inspection times.

constituent (Experiments 1-3) or the whole word (Experiments 4 & 5) frequencies were manipulated. Several frequency-driven parafoveal-on-foveal trends and effects we observed, but their expression was inconsistent. For example, in Experiments 1 and 3 only, the probability of skipping word n was greater if word n+1 contained a long low- compared to a long high-frequency initial constituent. Gaze duration also revealed frequency-driven parafoveal-on-foveal effects stemming from the initial constituent of the parafoveal word, but with an orthodox pattern arising in Experiment 2, an unorthodox pattern arising in Experiment 3, and with no significant differences in gaze in any of the other experiments.

Hyönä and Bertram used regression analyses to investigate the potential causes for their discrepant results. This revealed that foveal word length modulated the expression of the frequency-based parafoveal-on-foveal effects, with short foveal words tending to produce unorthodox parafoveal-on-foveal effects, while long foveal words tended to produce orthodox parafoveal-on-foveal effects. This trend was only true when the parafoveal word was short, suggesting that the word lengths of both foveal and parafoveal words play an important role in the expression of parafoveal-on-foveal effects.

Several corpus-based studies have been particularly successful in revealing frequency-based parafoveal-on-foveal effects. In one such study, Kliegl, Nuthmann and Engbert (2006) carried out regression analyses on the Potsdam Corpus, which revealed a small but significant frequency-driven parafoveal-on-foveal effect in single fixation duration. This 8ms effect, which was only reliable when the foveal word was short, fell in an orthodox direction. Since this effect was not mirrored in gaze duration, however, Rayner, Pollatsek, Drieghe, Slattery and Reichle (2007), have questioned whether it can truly be considered as representative of natural reading, especially since – as they argue – use of single fixation duration will bias towards the exclusion of certain word types (e.g., long low frequency words). Kennedy and Pynte (2005), however, obtained a similar pattern of results in gaze duration using the Dundee Corpus, in which 10 English and 10 French readers read 50,000 words taken from English and French newspapers. They observed a significant 12ms increase in gaze duration on short foveal words when the parafoveal word was low rather than high frequency; this parafoveal-on-foveal effect again fell in an orthodox direction and was not present for long foveal words. Subsequent analyses revealed that this effect was only significant for the English, but not the French readers[11].

It should be noted that drawing the strongest possible conclusion based on corpus studies such as these - that is, that parafoveal-on-foveal effects necessarily implicate parallel lexical processing – has not gone unchallenged. Most notably, Rayner et al (2007) raise particular issue with what they argue is

[11] The effect for the French readers was apparently sub-lexical in nature, with a sensitivity of the initial trigram of parafoveal words, but only when the foveal word was long.

a lack of control inherent in corpus studies. They suggest that low frequency words will often contain irregular letter strings, which as it has already been shown, can influence fixation duration on preceding words. Therefore, according to Rayner et al, Kliegl et al's frequency-driven parafoveal-on-foveal effect might simply have been an orthographic or sub-lexical parafoveal-on-foveal effect masquerading as one driven by frequency. In reply to Rayner et al's criticisms, Kliegl (2007) highlights the fact that even experimental studies are subject to problems of multicollinearity and further emphasises the point that evidence for parafoveal-on-foveal effects stemming from experimental studies are equally correlational in nature.

It is acknowledged that the research discussed here on lexical parafoveal-on-foveal effects has been presented with a degree of brevity. The purpose of this section, however, was not to provide an exhaustive commentary of these effects, but rather to highlight the main trends and controversies in the field: (a) like orthographic parafoveal-on-foveal effects, these lexical effects are also elusive, (b) their expression is often inconsistent, and (c) they are most consistently found in corpus analyses, which has led to some criticisms of these techniques.

2.2.3.2. Semantic Parafoveal-on-Foveal Effects: The Evidence

One means of determining whether parafoveal meaning can be extracted prior to fixation is to manipulate the pragmatic relationships between foveal and parafoveal words. In one such study, Murray and Rowan (1998) manipulated the relationship between an initial noun-phrase and a verb that followed, such that the verb was either plausible or implausible; for example, "*The hunters stacked*…" (plausible) versus "*The bishops stacked*…" (implausible). In their task, participants were instructed to read a sentence and then to press a button, the act of which triggered a second sentence to appear; the task being to decide whether the two sentences were physically identical or not; with data collected from just the first sentence analysed. Murray and Rowan observed increased first pass reading times on the noun phrase when its combination with the (yet to be fixated) verb was implausible compared to when the relationship was plausible. These results suggest that the plausibility of the verb had been detected prior to receiving a direct fixation. The degree to which this effect transfers to other reading tasks has, like Kennedy's experiments described above, been questioned, since it is argued that participants may have engaged task-dependent strategies (Rayner et al, 2003). Indeed, Rayner et al report a study carried out by Rayner and Miller that failed to replicate Murray and Rowan's effects using a more-usual reading task. Nevertheless, Murray and Rowan's results provide some suggestion that, albeit under certain conditions, the plausibility of a parafoveal word can be extracted before it is directly fixated.

Avoiding the criticisms associated with using so-called artificial laboratory tasks, Inhoff, Radach, Starr and Greenberg (2000) report obtaining high level parafoveal-on-foveal effects while engaging readers in 'natural' reading. Since no contingent changes took place in this experiment, its ecological validity can be considered as high as one can attain within an eye-tracking environment. In their study, participants were presented with sentences containing two critical words, the relationship between which was varied. There were three conditions, which were as follows: word n+1 was either (a) a repetition of word n (e.g., "mother's mother"), (b) a semantic associate of word n (e.g., "mother's father"), or (c) a word that was semantically unrelated to word n (e.g., "mother's garden"). Inhoff et al report an influence of word n+1 type on inspection of word n, with shorter gaze duration when word n+1 was either a repetition of word n, or a semantic associate, compared to when it was semantically unrelated (effect sizes: 26ms and 20ms, respectively). Furthermore, a separate analysis suggested that this effect cannot be attributed solely to cases in which foveal fixations fell within four character spaces of the parafoveal word, suggesting the results were not caused by oculomotor error. These results suggest that a semantic representation of a parafoveal word can be extracted while fixating the foveal word, which in turn influences foveal inspection times. Drieghe (2011), however, raises the possibility that these results may have been confounded with an increased tendency to skip word n+1 when the word pairs were either repeated or semantically related, compared to when they were semantically unrelated. Indeed, Inhoff et al do report a skipping pattern consistent with

Drieghe's observation. The validity of Drieghe's suggestion clearly requires further investigation, but this study remains one of the most convincing examples of semantic parafoveal-on-foveal effects reported to date.

Further evidence suggesting parafoveal-on-foveal meaning effects can be found with a more usual reading task comes from Rayner, Warren, Juhasz and Liversedge (2004), who manipulated the plausibility of a critical word, such that given the proceeding sentence context, the word was either (a) plausible, (b) implausible, or (c) anomalous. Pre-target gaze durations revealed a non-significant trend towards longer durations when the word to the right was anomalous compared to the other two conditions. Furthermore, when the dataset was restricted to just the fixations that were located within three character spaces of the target, the anomalous condition resulted in significantly higher gaze durations on the pre-target word compared to the plausible and implausible conditions, which did not differ from one another. Rayner et al suggest that the most parsimonious interpretation of these results is that they were caused by oculomotor error. That is, while the eye had intended to fixate the target word, oculomotor error caused it to fall short and to land on the pre-target word, from where word n+1 was processed. Whether semantic parafoveal-on-foveal effects such as these stem from attention being distributed to multiple words simultaneously or whether, as Rayner et al suggest, they are simply the result of a mislocated fixation, coupled with a stay-and-process strategy, remains highly controversial and is a topic that will be returned to in more depth in Chapters 5 and 6.

2.2.3.3. Higher Level Parafoveal-on-Foveal Effects: The Models

The SWIFT Model

As discussed above in the Section on ‘Low Level Parafoveal-on-Foveal Effects and the Models’ (2.2.2.2), the SWIFT model can account for parafoveal-on-foveal effects via the saccadic targeting mechanism. Specifically, since this mechanism is influenced by the relative activations of all words falling within the perceptual span, the orthographic and sub-lexical properties of a parafoveal word can influence foveal inspection times, and the same saccadic targeting mechanism can, qualitatively at least, account for the expression of frequency-driven and semantic parafoveal-on-foveal effects (Engbert & Kliegl, 2011; Schad & Engbert, 2012). A current challenge for the SWIFT model is whether or not is can simulate both orthodox (e.g., Hyönä & Bertram, 2004; Inhoff, Radach et al, 2000; Kliegl et al, 2006; Murray & Rowan, 1988; Rayner et al, 2004) and unorthodox (e.g., Hyönä & Bertram, 2004; Kennedy et al, 2002; 2004) lexical parafoveal-on-foveal effects. Given the dynamic and interactive nature of the SWIFT model, its prediction in this regard is currently unclear without dedicated simulations committed to addressing this question. Finally, while not modelled, it is clear that post-lexical parafoveal-on-foveal effects naturally fall out of parallel models, although this again remains to be simulated by the architects of SWIFT.

The E-Z Reader Model

Aside from Rayner et al's (2007) assertion that frequency-driven parafoveal-on-foveal effects, especially those found in corpus studies (e.g., Kliegl et al 2006), may result from potential positive correlations between a target word's normative word frequency and initial letter regularity, the E-Z Reader model – based on its core architecture – fails to account for any parafoveal-on-foveal effect past sub-lexical stages of word processing. In order for it to allow for lexical or post lexical parafoveal-on-foveal effects, the model must rely on their occurrence being caused by the consequences of a mislocated fixation. Specifically, oculomotor error would need to cause some saccades to fall short of their intended targets, with this coupled to a stay-and-process strategy, resulting in parafoveal words being processed from the suboptimal location of the foveal word. Research into this account is described in more detail in Chapter 6.

A great deal of theoretical importance has therefore been focussed upon whether mislocated fixations can account for parafoveal-on-foveal effects. Unfortunately, as Drieghe (2011) states "…the problem with the mislocated fixations account is, of course, that there is no way to experimentally determine whether a saccade has been mislocated or not" (p.

848). Another apparent limitation of the mislocated fixations account of lexical parafoveal-on-foveal effects is that it can only provide an account of orthodox parafoveal-on-foveal effects; predicting that the properties of the parafoveal word will be reflected on the foveal word in the same manner as if that word was directly fixated. Such an explanation clearly fails to provide an explanation for unorthodox parafoveal-on-foveal effects, where apparent processing difficulty of the parafoveal word, results in shorter foveal fixations. Given the theoretical importance placed on this interpretation in accounting for lexical parafoveal-on-foveal effects, it has become a timely imperative to investigate precisely how the oculomotor system responds in the event of a mislocated fixation. This is something I shall return to in Chapter 6.

2.2.4 Parafoveal-on-Foveal Effects: Summary

Orthographic and sub-lexical parafoveal-on-foveal effects are no longer considered controversial, with positive evidence now converging from multiple sources (e.g., Angele et al, 2011; Inhoff et al, 2008; Rayner, 1975; Starr and Inhoff, 2004). While it is suggested that these low level effects no longer differentiate between processing gradient and serial attention shift classes of model, this assertion would benefit from instantiation of the ‘pre-attentive’ low level attentional scan within the E-Z Reader model in order to demonstrate, quantitatively and not just qualitatively, the model’s capability in simulating these effects. Lexical parafoveal-on-foveal effects, however, remain a fiercely debated topic. Not only is their occurrence highly contentious (e.g., Kliegl et al, 2006; Kliegl, 2007; Rayner et al, 2007; 2003), the plausibility of the argument that they can be accounted for by a mislocated fixation followed by a stay-and-process response within a serial architecture is also hotly debated (e.g., Drieghe et al, 2008; Drieghe, 2011; Engbert, & Kliegl, 2011; Kennedy, 2008). A means of measuring the response to a mislocated fixation is therefore required before any strong statement can be made regarding the potential implications the existence of higher level parafoveal-on-foveal effects might have for models of eye movement control. As stated above, this topic will be returned to in Chapter 6.

2.3 Word n+2 Preview Benefit

An alternative route for investigating the link between fixation location and attention during reading was introduced by McDonald (2005). He reasoned that, according to the Processing Gradient framework, all words falling within the effective span of apprehension – including word n+2 - should benefit from an accumulative preview benefit and this should be reflected in the eye movement record. While acuity constraints dictate that the size of the benefit will be smaller for words in position n+2 compared to n+1, if words are processed in a parallel fashion, then a benefit should nevertheless be present.

2.3.1 Word n+2 Preview Benefit: The Models

The SWIFT model

Word n+2 preview benefits are operationalized within the SWIFT model via several mechanisms: the targeting mechanism, the foveal inhibition mechanism and the zoom lens mechanism (Schad & Engbert, 2012; Risse, Hohenstein, Kliegl & Engbert, 2014). Specifically, since the model allows – under optimal conditions - for multiple words falling within the perceptual span to be lexically processed simultaneously, word n+2 can be pre-processed while fixating word n. The consequences of this are twofold. First, pre-processing influences lexical activation levels, which in turn influences which word is targeted for an upcoming fixation. Second, upon fixation, word n+2 will require more processing if parafoveal preview had previously been denied, increasing the probability that the foveal inhibition mechanism will delay the onset of a saccade out of the word, which will increase inspection times (Engbert & Kliegl, 2011).

The presence and/or strength of word n+2 preview effects are further modulated by the 'zoom lens' mechanism (Schad & Engbert, 2012). As will be recalled from Chapter 1, this mechanism allows the processing span to dynamically increase once the foveal word enters its lexical completion state. It is within this stage of decreasing foveal activity that a narrowly focussed lens will begin to expand, and in so doing increase the possibility that word n+2 will be pre-processed. Since activation levels are negatively correlated with a word's frequency in SWIFT, a high frequency foveal word will – all things being equal - enter its lexical completion state sooner and as such will benefit from an extended processing span more than low frequency foveal words. For word n+2 preview benefits to optimally occur, therefore, word n should be an 'easy' to process word.

The E-Z Reader Model

The E-Z Reader model (Reichle et al, 2009) can also allow for word n+2 preview benefit providing a series of sequential processes complete within the limited timeframe imposed by saccadic latency. Specifically, the command to plan a saccade out of word n is tied to the completion of L1 on word n; the model specifies that this saccade will be committed to execution within, on average, 100ms and executed 25ms thereafter. This saccade can be cancelled and a new one programmed to word n+2, providing stage L1 completes on word n+1 while the saccadic mechanism is still in its labile stage. This cancellation will incur a small time penalty. Word n+2 preview benefit can thus be accommodated within the model providing all the following conditions are satisfied within 125ms (or marginally longer if word n+1 is skipped): the L2 stage of word n must complete allowing attention to shift to word n+1; both the L1 and L2 stages of processing on word n+1 must complete allowing attention to precede once more onto word n+2; word n+2 must then undergo some degree of pre-processing. While this scenario is not outwith the bounds of possibility, it is likely to constitute an exceptional set of circumstances, especially since the model specifies that each attention shift will take an average of 50ms (Reichle et al, 2009)[12].

2.3.2. Word N+2 Preview Benefit: The Evidence

The first study to investigate word n+2 preview benefit was conducted by McDonald (2005). He selected two to four spatially adjacent words from the Dundee English Corpus (Kennedy, 2003), which between them contained a sequence of three progressive saccades; care was taken to only include those cases in which the distance between the first and third (target) fixation spanned no more than 15 character spaces. Contrary to the predictions of Processing Gradient models, McDonald found no evidence for a cumulative preview benefit regardless of whether (a) the eccentricity of preceding fixations, or (b) the summed duration of preceding fixations were included as predictor variables for first fixation duration on the target word.

Following McDonald (2005), Rayner Juhasz and Brown (2007) tested the presence of word n+2 preview benefit experimentally. To accomplish this, they used the boundary paradigm but in a novel way: rather than placing the boundary location on the space prior to the target word, they placed it on the

[12] Recent simulations indicate that the E-Z Reader model can, via the mechanism just outlined, accommodate word n+2 preview benefits that are orthographic in nature (Schotter et al, 2014). The results of this simulation will be returned to in the Discussions of Chapters 4 and 5.

space prior to the *pre-target* word. Therefore, any observed differences on the target word n+2 must originate from a word that falls two words - or potentially more depending on whether word n was fixated - upstream. As can be seen in the example below, Rayner et al used either an identical, alternative or nonword preview. The vertical line represents the position of the invisible boundary when word n+2 preview benefits were tested. They also included a condition in which the boundary was placed immediately before the target word, thereby testing word n+1 preview benefit in the traditional manner. To be clear, the same set of target words were used when testing both word n+1 and word n+2 preview effects; with just the position of the boundary dictating which of these was being tested. While their results revealed the standard word n+1 preview benefit, readers appeared to be immune to the word n+2 preview manipulation. There were no word n+2 preview effects on the subsequent inspection of either the combined pre-target region (e.g., “the large”), or the target word region (e.g., “carrots”).

Example word n+2 preview manipulation prior to passing the boundary:

```
Identical:                    *
John used a knife to chop the | large carrots for dinner last night.
Alternative:
John used a knife to chop the | large allergy for dinner last night.
Nonword:
John used a knife to chop the | large xonnulc for dinner last night.
```

Example word n+2 preview manipulation after passing the boundary:

```
Identical:                         *
John used a knife to chop the | large carrots for dinner last night.
```

To optimize conditions for parallel lexical processing, Rayner et al conducted a second experiment with modified materials. To encourage the target word to fall within the effective span of apprehension, the pre-target and target words were 3 to 4 characters in length, for example, “cozy aura”. They also assumed the roles of adjective-noun, under the assumption that the adjective can only be fully interpreted in conjunction with the noun. Despite these modifications, Rayner et al drew the same conclusions from the results of this second experiment as they had their first: although they obtained a word n+1 preview benefit, there was no evidence for a word n+2 preview benefit, and therefore no evidence that word n+2 had been pre-processed.

However, from inspection of Rayner et al's materials, it can be seen that in many cases it was a relatively low frequency word in the position of word n+1. While the mechanism modulating the size of the perceptual span in the SWIFT model – the zoom lens – is only influenced by foveal input, it is feasible that a low frequency intervening word may have constrained attentional resources. Also, across both experiments, word n was almost always a short, high frequency function word. Since this word type is skipped approximately two-thirds of the time of during reading (Rayner, 2009), the final fixation prior to passing the boundary is likely to have frequently originated from a location three words upstream from the target, and neither class of model would be likely to predict a word n+3 preview benefit.

Consequently, Kliegl, Risse & Laubrock (2007) replicated Rayner et al's experiment but employed longer high frequency words in the position of word n (average: 7 characters) and 3-letter high frequency function or content words in the position of word n+1. The word n+2 previews were either identical or alternative nonwords. Kliegl et al obtained several effects indicative of parallel lexical processing. The first relates to word n+1 type: gaze duration was 26ms longer when the word to the right contained a content compared to a function word. As discussed in the preceding section, these apparent lexical parafoveal-on-foveal effects are difficult to reconcile with a model that assumes lexical processing advances in a strictly serial sequential manner; and this topic will be returned to in more depth in Chapter 6.

Consistent with Rayner et al’s results, there was no effect of word n+2 preview on word n+2. However, there was an 11ms increase in gaze duration on word n+1 when word n+2 had received an invalid preview and a 15ms increase in gaze duration on word n, although this latter effect was only present when word n+1 was a content word. Kliegl et al interpreted these numerically small but significant results on words n and n+1 as general evidence supporting distributed lexical processing.

It has been argued, however, that Kliegl et al’s results can be explained within the architecture of the E-Z Reader model if one factors mislocated fixations into the model (Schotter, Angele & Rayner, 2012). For example, the effects of word n+1 type and word n+2 preview on word n can be explained if one assumes that these words were targeted, but due to saccadic error the saccade undershot its target, and this was associated with a stay and process response. Such an explanation would, however, necessitate a fixation mislocated by 4 or more character spaces in order to account for the expression of the word n+2 preview effect on word n. But while a mislocation of this magnitude seems improbable, it cannot be entirely discounted as a potential explanation for Kliegl et al’s result.

Regarding the expression of word n+2 preview effects on word n+1, it could be suggested that word n+1 was identified parafoveally, resulting in the cancellation of a saccade to word n+1 and the initiation of a new saccadic program to word n+2; in the extra time required to plan and execute the saccade to word n+2, a word n+2 preview benefit is obtained, but because of saccadic error, the saccade then undershoots word n+2 and lands on word n+1. Rather than relocating to word n+2, it could be argued that word n+2 is then processed from this suboptimal position, reflecting word n+2 preview benefit on word n+1. In agreement with this suggestion, in a follow-up study using identical procedures and materials, Risse and Kliegl (2011) not only replicate their previous word n+2 preview effect on word n+1, but also uncovered an 18ms preview benefit on word n+2, but only in those cases in which word n+1 had been skipped. While the plausibility of the mislocated fixations hypothesis will be returned to in Chapter 6, a further adaptation of the word n+2 boundary paradigm does help to adjudicate between whether or not a planned skip of word n+1 could be responsible for these effects.

Based on the above logic, the E-Z Reader model should not predict a word n+2 preview benefit if word n+1 preview is also denied in parafoveal vision, since the condition that word n+1 has been processed - allowing attention to legitimately shift to word n+2 - cannot then be satisfied. Radach, Glover and Vorstius (2007) employed a similar design and used similar materials to Kliegl et al; but manipulated the previews of both words n+1 and n+2 such that prior to passing the invisible boundary located just before word n+1, one, the other, neither or both of words n+1 and n+2 contained a nonword letter string preview. In addition to obtaining a standard word n+1 preview benefit, Radach et al also obtained effects of word n+2 preview on word n+1 when word n+1 had also received an invalid preview, suggesting that these effects could not have stemmed from a failed skip of an identified word n+1. Importantly, gaze duration also revealed a significant word n+2 preview benefit on word n+2 when word n+1 had also received an invalid preview. These results suggest that word n+2 preview benefits cannot be explained by a failed skip of an identified word n+1 coupled with a stay and process response to a mislocated fixation.

A potential explanation for why Risse and Kliegl (2011) only obtained word n+2 preview benefit when word n+1 had been skipped could be related to the fact that once the eye passed the boundary, word n+2 was always available for pre-processing. Therefore, for cases where word n+1 was fixated, a preview benefit for word n+2 could start accruing during word n+1 fixation, diluting any word n+2 preview benefit originating from word n.

As an interim summary, the evidence discussed so far would seem to indicate that word n+2 preview benefit does exist, but only when word n+1 is short and all words within the effective span of apprehension are high frequency.

To investigate whether the properties of word n+1 modulate word n+2 pre-processing, Angele, Slattery, Yang, Kliegl & Rayner (2008) conducted an invisible boundary experiment similar to that of Radach et al, in which one, the other, neither or both of words n+1 and n+2 received an invalid preview prior to passing an invisible boundary located before word n+1. Word n+1 was either a high or low frequency word. Unlike Radach et al (2007), however, they obtained no evidence that word n+2 preview influenced fixation durations on words n, n+1 or n+2, irrespective of word n+1 frequency. This null result could however be linked to Angele et al's choice of stimuli. While both words n and n+2 were of a high frequency (177 and 175 per million, respectively) the lengths of these words ranged from 3 to 13 characters (mean 7), and word n+1 ranged from 4 to 10 characters (mean 6). On some occasions, therefore, word n may not have attracted a fixation, while on others, even if word n was fixated, word n+2 may have fallen outwith the effective span of apprehension. Indeed, the authors acknowledge that word n+1 length may have been a contributing factor in their null result, concluding that word n+2 preview benefits may only be restricted to cases in which word n+1 contains just three characters.

A follow-up article published by Angele and Rayner (2011), which reports two separate experiments, does however provide further insights into the nature of word n+2 preview benefit. In Experiment 1, they employed the same preview manipulations as Angele et al (2008) (minus the condition in which a nonword n+2 preview followed an identical word n+1 preview), and similar to Kliegl et al (2007), they manipulated word n+1 type such that it was either a high frequency content word or an article. According to Radach's word grouping hypothesis (1996, cited in Drieghe, Pollatsek, Staub & Rayner, 2008) article-noun word pairs are processed as one perceptual unit and so should provide the optimal conditions for word n+2 pre-processing (cf. Drieghe et al, 2008). The three important results relevant to the current discussions are: First, in line with the above-mentioned research on orthographic parafoveal-on-foveal effects, gaze duration and go-past time revealed an orthographic parafoveal-on-foveal effect of word n+1 preview on word n. Second, when word n+1 had been unavailable, a nonword preview of word n+2 appears to have attracted attention directly towards it, reflected in a higher skipping probability for word n+1 when word n+2 had received a nonword preview. Finally, unlike Kliegl et al (2007), they obtained no evidence for a lexical parafoveal-on-foveal effect of word n+1 type on word n, or word n+2 preview on words n or n+1 inspection times.

These results suggest that orthographic irregularity can be identified parafoveally in word n+1 *and*, potentially, word n+2 when word n+1 is also orthographically illegal. Angele and Rayner suggest, however, that because these effects are not lexical in nature, they can be explained within the architecture of the E-Z Reader model via the low level attentional scan detecting the upcoming irregularity. As discussed in Chapter 1, proponents of the model have previously suggested that orthographic parafoveal-on-foveal effects can be explained in their model via such a mechanism (e.g., Rayner et al, 2007), although whether the same mechanism can plausibly account for an apparent word n+2 attraction in the case of irregular previews remains unclear. Indeed, suggesting that a parafoveal word (n+1) can be skipped prior to its lexical recognition because a remote parafoveal word (n+2) contains irregular properties seems somewhat incongruous for a model postulating that lexical access is the engine that drives the eyes through text.

In their second experiment, Angele and Rayner (2011) tested whether word n frequency modulates word n+2 preview benefit. Preview conditions were exactly the same as Angele et al (2008), and word n+1 always contained the article “the”; word n frequency was either high or low. Once again they obtained an orthographic parafoveal-on-foveal effect with longer gaze and go-past time on word n when the word to its immediate right was a nonword. They also obtained a word n+2 preview benefit that was reflected on both words n+1 (go-past time) and n+2 (gaze durations), but only in cases where word n+1 had also received an identical preview. So, unlike Radach et al (2007), Angele and Rayner did obtain word n+2 preview effects, but consistent with the E-Z Reader model, only in those cases in which a failed skip of word n+1 could have been implicated. Indeed, word n+2 preview benefit only achieved significance on word n+2 when word n+1 had also been skipped. Finally, they also found that when word n was a high frequency word, an invalid word n+2 preview resulted in initial fixations that landed closer to the beginning of word n+1. Therefore, unlike their first experiment in which invalid previews of words n+1 and n+2 resulted in an apparent attraction towards word n+2, their second experiment suggested that an invalid preview of word n+2 can result in what could be interpreted as a more cautious reading style, with earlier landing positions on word n+1.

It seems therefore that word n+2 preview benefits are incredibly sensitive, with subtle differences between item sets potentially triggering differing responses. Whether different targeting responses can be accounted for within the E-Z Reader model remains unclear, but given the dissociation between serial lexical processing and low-level pre-attentional scans, it is a possibility. An attempt to implement the low level scan in the E-Z Reader model, together with simulations modelling how such targeting modulations can occur would be invaluable for understanding whether the E-Z Reader model can cope with this patterns of effects.

Given that Rayner et al (2007) and Angele et al (2008) both failed to obtain word n+2 preview benefits, while Kliegl et al (2007), Risse and Kliegl (2013), Radach et al (2007) and Angele and Rayner (2011) all did, strongly suggests that word n+1 length and frequency are important factors. As detailed previously, word n length may also have played a critical role in the null effects reported by both Rayner et al (2007) and Angele et al (2008). Therefore, testing the range over which these effects can occur while keeping word n of medium length and high frequency will be important when attempting to dissociate the factors that potentially modulate the expression of word n+2 preview benefit.

One study that has approached word n+2 preview benefits from a slightly different angle is that of Radach, Inhoff, Glover and Vorstius (2013). They created a word n+2 preview benefit experiment using the boundary paradigm, manipulating just word n+2 (not n+1). They presented participants with sentences containing either a high or low predictable word in the position of word n+2, for example: "*Ashley quickly vacuumed the carpet/stairs before her friends arrived for the party*". Word n+2 received one of three previews: (a) identical (no change), (b) alternative word (e.g., "stairs" would act as the preview for "carpet" and vice versa), or (c) a nonword preview (e.g., "cwoyok"). Like Angele et al (2011; Exp1), they found that a nonword preview of word n+2 appeared to attract attention directly towards it, with a higher probability of skipping word n+1 in those cases. Furthermore, when word n+2 was highly constrained given the preceding sentence context, gaze duration was 20ms shorter on word n+2 when it had received a valid preview compared to the other two conditions combined.

Importantly, this effect remained when the analysis was restricted to cases in which word n+1 had also been fixated prior to fixating word n+2, suggesting that word n+1 had not been identified parafoveally. This subset of data revealed another significant effect in gaze duration on word n+2:

specifically, a high predictable n+2 preview changing to a low predictable n+2 target significantly increased fixation durations compared to the other two other preview types. This result is interesting as it suggests that the highly predictable word n+2 had been processed while fixating word n and when that input changed to a word that was low in predictability, the reading process was disrupted. This may not have occurred when n+2 was previewed as a word low in predictability, since the processing of that word might not have been as advanced (due to its low predictability) and consequently it was not associated with the same level of disruption when the input changed. This result is significant since it provides the first piece of evidence pertaining to a word n+2 preview benefit that is lexical in nature.

Finally, Radach et al (2013) report obtaining a near-significant (10ms; $p<.06$) spillover effect in gaze duration of word n+2 preview on word n+3, in which an identical preview resulted in shorter fixation durations on the subsequent word. This effect also appears to be in conflict with the predictions of the E-Z Reader model, as the trigger to program a saccade out of a word is the completion of L1 on that word, which requires the extraction of the orthographic code. Therefore, effects of preview should always be confined to the word from which they originated (or earlier in the case of parafoveal-on-foveal effects) and never spillover.

2.3.3 Word n+2 Preview Benefit Summary

Research into word n+2 preview benefits is bedevilled with inconsistent results. Several themes are however forming in the literature. First, readers often demonstrate sensitivity to word n+2 previews, although the expression of this sensitivity manifests in differing ways in different experiments. For example, some studies report obtaining effects in durational measures on words n+1 and/or n+2 (e.g., Angele & Rayner, 2011, Exp2; Kliegl et al, 2007; Radach et al, 2007; Radach et al, 2013; Risse & Kliegl, 2011), while in others, targeting decisions are influenced (e.g., Angele & Rayner, 2011; Pynte et al, 2004; Radach et al, 2013). Second, the importance of creating optimal conditions for parafoveal processing is now clear, with the positive results restricted to those cases in which word n+1 contained just three characters (Angele & Rayner, 2011; Kliegl et al, 2007; Radach et al, 2007; Radach et al, 2013; Risse & Kliegl, 2011). Third, the majority of positive effects fail to distinguish between the E-Z Reader and SWIFT models since their expression is often consistent with a double attention shift followed by a mislocated fixation and stay and process response, making them - in theory at least - compatible with the E-Z Reader model.

It seems that in order for word n+2 preview effects to differentiate between whether attention is distributed serially or in a parallel fashion during reading, several questions need to be addressed. First, the only evidence pertaining to a word n+2 preview benefit using the English writing system has, to date, been restricted to cases where word n+1 was short and not fixated (Angele & Rayner, 2011); the question therefore arises as to whether these effects can be obtained following a fixation on word n+1 when greater control over words n and n+2 length is exercised. Second, can these effects be obtained over longer ranges? If they can, then that would reduce the probably that (a) word n+2 preview effects are the result of a double attention shift, and (b) word n+2 preview effects expressed on word n+1 are the result of a failed skip of word n+1 followed by a stay and process response. Third, given the evidence for semantic preview effects (Hohenstein et al, 2013; Schotter, 2013), might these effects also exist if the manipulation occurs on word n+2? Providing evidence for advanced lexical processing of a word two words downstream would provide a stronger foundation on which the E-Z Reader and SWIFT models could more easily be differentiated.

All these questions remain to be determined via controlled experiments. One final question rests of the shoulders of model architects. Risse, Hohenstein, Kliegl and Engbert (2014) have provided quantitative simulations demonstrating SWIFT's capabilities in reproducing word n+2 preview effects on word n+1[13]. It would similarly be an invaluable modelling exercise for proponents of the E-Z Reader model to demonstrate,

---

[13] The boundary paradigm was simulated by attributing the nonword preview the lowest frequency and resetting the activation level upon fixation. It should be noted that while effects of word n+2 preview on words n and n+1 were simulated reasonably well, those on word n+2 were overestimated.

quantitatively and not just qualitatively, the extent of the model's capabilities in this regard.

2.4. Future Directions

The possibility has recently been raised that research into the serial versus parallel debate may have reached an impasse, or at the very least, that the predictions derived from serial and parallel models have become too closely aligned to allow strong conclusions to be drawn regarding how attention might be distributed during reading (Murray, Fischer & Tatler, 2013). Murray et al suggest that the debate could continue via the investigation of how attention is distributed across phrases (e.g., Radach, 1996), rather than concentrating on individual word units. But there also appears to be some ways in which the predictions of serial and parallel models still might be teased apart within the context of the research discussed in this chapter. Some of these will be investigated in the following chapters, but first one of the underlying assumptions of the parallel perspective will be investigated.

CHAPTER 3

Can Isolated Word Pairs be Lexically Processed in Parallel?

3.1. Introduction

Since the introduction of quantitative models of eye movement control (e.g., E-Z Reader: Reichle et al, 1998; SWIFT: Kliegl & Engbert et al, 2003 and Glenmore: Reilly & Radach, 2006), there has been a sustained drive in research targeted at determining whether multiple words can be lexically processed in an overlapping fashion[14]. It will be recalled from Chapter 2 that the majority of this research has centred on identifying effects considered incompatible with the serial perspective, such as semantic preview benefit, word n+2 preview benefit and parafoveal-on-foveal effects. Since the late nineties, however, the E-Z Reader model – currently version 10 – has successfully evolved to account for many of the effects once thought to be a challenge to it; rendering some of these effects less diagnostic than originally thought. As the findings that serial and parallel models can account for become increasingly aligned, the question arises as to whether it is, at the very least, psychologically plausible to assume that two words can be lexically processed in a parallel fashion. Indeed Reichle, Liversedge, Pollatsek and Rayner (2009) recently stated “we know of no compelling empirical evidence supporting the claim that two or more words

[14] Indeed, an analysis by Murray et al (2013) shows that the number of published papers referring to serial and parallel processing within the context of reading had risen from just 8 citations/year in 1997 to 108 citations/year in 2012.

can be simultaneously processed" (pp116). This study therefore sought to investigate this question.

Both semantic preview benefit (e.g., Hohenstein & Kliegl, 2013; Schotter, 2013), and lexical and post–lexical parafoveal-on-foveal effects (e.g., Hyönä & Bertram, 2004; Murray & Rowan, 1998) suggest that words are being lexically processed in a parallel fashion. These effects are, however, controversial, since they might have resulted from a variety of artefacts unrelated to parallel lexical processing, such as mislocated fixations (see Chapter 2) or calibration error (e.g. Drieghe, 2011; Reichle & Drieghe, 2015).

While parallel lexical processing remains contentious, there is some evidence that, at the very least, words can be lexically processed in the parafovea, evidenced by the increased probability of skipping high compared to low frequency words (e.g., Angele et al, 2008; Henderson & Ferreira, 1993; Rayner & Fischer, 1996; Rayner, Sereno & Raney, 1996). While these skipping effects are confounded by the possibility that high frequency words might simply be 'guessed' more often than low frequency words, resulting in the higher skipping rates for this class of words[15], it is generally considered uncontroversial that parafoveal word identification contributes towards word skipping (Rayner et al, 2003). Taken together, a variety of effects obtained from reading studies strongly suggest that words can be lexically processed in

[15] Indeed, there is some evidence that parafoveal words are not always parafoveally identified prior to being skipped (e.g. Balota et al, 1985; Drieghe, Rayner & Pollatsek, 2005).

the parafovea and, potentially, also in parallel with the foveal word, although there are a variety of alternative explanations for some of the seemingly-parallel effects.

Results from methodologies that, arguably, might not engage participants in natural reading behaviours have long been used as a means for investigating the underlying processes inherent in reading. For example, Howes and Solomon (1951) and Whaley (1978) reported faster and/or more accurate responses to high frequency words compared to low frequency words in tachistoscopic identification[16] and lexical decision tasks[17], respectively. Additionally, Meyer and Schvaneveldt (1971) also demonstrated, via a lexical decision task (Exp 1) and a same-different matching task[18] (Exp 2), a response time advantage for simultaneously presented semantic associates compared with unrelated word pairs. In another classic study, Rubenstein, Lewis and Rubenstein (1971) showed, again via a lexical decision task, that pseudohomophones were responded to slower than alternative pronounceable nonwords, suggesting that pseudohomophones had accessed the associated word's lexical representation. But while these studies do not engage participants in 'natural' reading tasks, there is evidence that such

[16] Tachistoscopic identification tasks require participants to verbalise a briefly presented stimuli.
[17] Lexical decision tasks require participants to decide whether a word is a word or a nonword as quickly and as accurately as possible.
[18] Same-different matching tasks require participants to decide as quickly and as accurately as possible whether two letter strings (or sentences) are physically identical or not.

results are also reflected in the eye movement measures reported in eye tracking studies during more natural reading (e.g., Milliet & Sparrow; 2004; Schilling, Rayner & Chumbley, 1998; Schotter, 2013). Tasks involving isolated words or word pairs, such as these, have thus provided useful ‘baseline data’ upon which theories of lexical access (e.g., Forster, 1976; Morton, 1969; Coltheart et al, 2001) and models of eye movement control (Reichle et al, 2009; Schad & Engbert, 2012) have been formulated.

It appears, therefore, that such tasks have the potential to inform on whether parallel lexical processing is possible. While it has been suggested that tasks involving isolated words, such as those just mentioned, lack the ecological validity of eye tracking experiments (e.g., Rayner et al, 2003; Schotter et al, 2012), even eye-tracking experiments have their limits in terms of ecological validity. For example, preventing head movements, the use of bite bars, being monitored and answering comprehension questions are all likely to cause readers to adjust their ‘natural’ reading habits. Given the clear correlates between the results of eye tracking tasks involving sentence reading and so-called ‘non-reading’ tasks, such as isolated word presentation (Schilling et al, 1998), the utility of the latter should not be underestimated. Indeed, they may offer an interesting route into investigating how attention is distributed during reading, a feat that, to date, eye-tracking studies have failed to access.

A variety of studies dating back to the 1970's have used 'non-reading' tasks to investigate foveal and parafoveal word identification. In an early example of such work, Rayner and colleagues (Rayner, 1978; Rayner, McConkie & Enrlich, 1978; Rayner McConkie & Zola, 1980) conducted a series of experiments that sought to investigate the nature of parafoveal processing. Typically they asked participants to fixate a central cross; and once a fixation was detected, a word was presented at 1, 2 or 5 degrees of visual angle either to the left or to the right of fixation. Participants were asked to orientate toward this word as rapidly as possible and to name it. Analogously with the boundary paradigm (Rayner, 1975), various properties of the parafoveal word changed during the saccade, meaning that the form of the word to be verbalised differed from the word seen parafoveally. This paradigm allowed Rayner et al to test the limits over which parafoveal processing can occur and the properties that did and did not impinge on parafoveal processing. Interestingly, this task appears to have engaged a similar asymmetric span of attention as is found in reading studies (e.g., McConkie & Rayner, 1976), with shorter naming latencies when the parafoveal word fell to the right, and with increased latencies with increasing eccentricity. The time available for parafoveal processing was necessarily restricted to saccadic latency, which with these experimental paradigms, typically fell between 160ms and 200ms. Within these tight time constraints, it was apparent that abstract letter codes could be extracted from the parafovea, with the first two characters of the word providing the most facilitation. These durations did not, however, appear to permit the extraction of the semantic code of the parafoveal word.

When a parafoveal word is available for longer durations, there is some suggestion that higher level processing may occur on that word. For example, Balota and Rayner (1983) found that when a parafoveal nonword was accompanied by the simultaneous presentation of a foveal row of xxx's, saccade latencies increased to an average of 255ms. As before, participants were asked to name the parafoveal word, which changed as a saccade brought that word into foveal view. There was a small but significant 7ms reduction in pronunciation latency when the peripheral nonword was orthographically related to a semantic relative of the target, compared to when there was no semantic relation. For example, 'snckks' appears to have primed 'snakes', which resulted in shorter naming latencies for the semantically related 'lizard' compared to the semantically unrelated 'limits'. This effect was only significant when the parafoveal word fell to the right, but was replicated in a second experiment. Balota and Rayner did not report analyses by-items and suggest that this effect is most likely due to the particular allocation of different items to each condition. This interpretation was encouraged by the lack of evidence for semantic preview benefit at the time the paper was written.

It is also however possible, that when saccade latencies are increased, and there is longer parafoveal preview, that some lexical processing can be performed on the parafoveal word and if this is sufficiently advanced it might lead to semantic priming. More recently, Baccino and Manunta (2005) and Simola, Holmqvist and Lindgren (2009) have looked at whether semantic information can be extracted from parafoveal words using eye-fixation related potentials (EFRPs). In both experiments, participants were presented with a central fixation cross followed by the simultaneous presentation of a central and a parafoveal word. In the Baccino and Manuta study, the parafoveal word always fell to the right of the central word, while in the Simola et al study it fell either to the left or to the right. In both cases, participants were asked to read both words silently and then to fixate a second cross that was always located at a point beyond the peripheral word. When this cross was fixated, participants were required to perform a semantic judgement task. In both experiments, there were three relations between the central and peripheral words, they were: semantic associates, non-associates, or the peripheral target was a nonword. While Simola et al observed shorter first fixations on the central word when the peripheral target was a nonword falling to the right, neither experiment provided evidence for a semantic parafoveal-on-foveal effect on inspection durations. The ERP analyses, however, (which were restricted to fixations on the central word), in both experiments showed differences in activity reflecting an early detection of the peripheral nonword. But only Baccino and Manunta uncovered evidence suggesting the extraction of parafoveal semantic information, finding that semantic associates could be differentiated from non-associates via a P2 component that peaked at 215ms within the frontal to occipital regions. Simola et al suggest that their failure to replicate this semantic effect may have been due to task differences, with their task requiring more time dedicated to planning and executing a saccade to the peripheral word as a result of the variability of its location, which was not the case in the Baccino and Manunta study. Again, these experiments strongly suggest that orthographic properties of parafoveal words can be extracted in the parafovea, and again, there is an indication that semantic information can be extracted from a parafoveal word, albeit under favourable conditions.

All studies discussed thus far have focussed on either the sub-lexical (i.e., orthographic) or post-lexical (i.e., semantic) properties of parafoveal words, but nothing has been said regarding lexical factors. This is important since it is generally agreed that the process of lexical identification occurs in stages. Specifically, for each word, the first stage in its identification involves the extraction of orthographic information from the page. Only once this information has been retrieved can the lexical characteristics of a word influence ease of processing and therefore affect identification time. Lexical factors are factors linked to, or dependent on, the process of word recognition, such as a word's frequency, phonology or morphology. Most theories of word recognition suggest that only after these characteristics have influenced the word recognition process, with the relevant lexical entry identified, can the semantic properties of a word then be extracted. It therefore follows that while positive evidence for semantic information extraction implies lexical processing has taken place, null reports cannot determine that lexical processing did not occur - only that there was insufficient time to complete the later process of semantic extraction. Equally, since orthographic processing is assumed to precede lexical processing, reports of orthographic effects also fail to capture the extent of lexical processing. Consequently, if we wish to determine the extent of lexical processing taking place on a parafoveal word, we must investigate a lexical characteristic that has been shown to affect the speed with which a word is lexically identified, such as a word's frequency.

Effects of word frequency are well documented in both isolated and multiple unconnected word tasks (e.g., Chambers & Forster, 1975; Howes & Solomon, 1951; Murray & Forster, 2008; Schilling et al, 1998; Whaley, 1978) and in sentence reading tasks, where frequency effects are often realised in the duration of the very first fixation falling on a word (e.g., Ashby et al, 2004; Rayner & Duffy, 1986; Schilling et al, 1998). Given the robust nature of the frequency effect and its early impact on the eye movement record, it is clear that this lexical characteristic is well suited to testing the existence of early parafoveal lexical processing.

While some studies, similar to those described above, have systematically manipulated lexical characteristics, such as parafoveal word frequency, they tend not to do so in isolation making it difficult to determine the origin of any effects. For example, Vitu, Brysbaert and Lancelin (2004) asked participants to read three simultaneously presented letter strings from left to right, the first and middle letter strings both formed words of either a high (left) or low (middle) frequency, or vice versa, while the letter string on the right was always a row of xxx's. The two words were either orthographic neighbours (e.g., pour/four) or not (e.g. pour/clan). Fixation duration on the left (foveal) word was measured as a function of the word pair's frequency relationship, their orthographic relationship, and whether the letter substitution for orthographically similar word pairs occurred internally or externally. Vitu et al found that gaze duration was shorter when the parafoveal word was an orthographic neighbour, but that this was modulated by (a) the frequency relationship between the two words, and (b) whether the substitution occurred internally or externally. Thus, while the lexical characteristic of word frequency was manipulated in the parafoveal word, so was the frequency of the foveal word, making it difficult to determine which of the two was driving the observed patterns of effects.

To remedy this apparent gap in the literature, this study investigated whether effects of parafoveal word frequency can be obtained when the task requires the simultaneous lexical processing of a foveal and a parafoveal word. A task was chosen that (a) was capable of tapping into the lexical processing of two simultaneously presented words, and (b) could be run under conditions that would necessitate those words be processed in a parallel fashion: same-different word matching.

It has long been known (e.g. Chambers & Forster, 1975) that when two vertically aligned letter strings are presented simultaneously, the time to decide whether they are physically identical or not follows, at least for the ‘same’ decisions, the following pattern: the fastest and most accurate responses are associated with high frequency words, followed by low frequency words, and then by legal nonwords, with illegal nonwords associated with the slowest response times and highest error rates. It therefore appears that although the task could potentially be carried out purely on the basis of comparisons related to physical identity, derived from the letter shapes, for ‘same’ decisions at least, other lexical and/or sub-lexical factors exert an effect on the time taken. If, however, there is a mismatch between the two letter strings and a ‘different’ decision is required, such effects of lexical status and word frequency typically disappear.

To account for this pattern of findings, Chambers and Forster proposed a multi-level ‘race’ model that suggests that comparisons between the two letter strings proceed simultaneously at letter, letter cluster and lexical levels of analysis. Whichever level of analysis returns an answer first wins the ‘race’ and determines reaction time. According to this theory, the time required to make a decision at each level of analysis will depend upon both the time taken to derive a representation at that level and on the number of comparisons required to make the judgement. Therefore, within the model, a trade-off exists where more time is required to derive the higher level representations, but less time is then required to complete the comparison process involving higher level representations, since fewer comparisons are necessary.

To elaborate, in the case where both strings of letters form words, lexical information can be retrieved and one comparison made; familiar letter clusters can also be grouped and compared as several ‘chunks’; while strings of unrelated letters must be compared on an individual letter-by-letter basis. So, when the two letter strings comprise the same word, the lexical representations for those words can be sought and just one comparison made, resulting in this level to winning the ‘race’ and this is reflected in the expression of a frequency effect. Legal nonwords on the other hand will require at least two (or more) comparisons to be completed at a letter cluster level, which, assuming that comparisons involve a significant amount of time, will inflate response times compared to the case when only a single comparison is required, thus, producing an effect of lexicality. Finally, despite the potentially faster extraction of single letter level representations compared to letter clusters the latter will require fewer comparisons, potentially allowing the letter cluster level to win the ‘race’, resulting in an effect of orthography.

In contrast to ‘same’ response times, ‘different’ response times have usually been found to not usually exhibit any lexical effects, but sometimes with effects of orthography (i.e., faster response times for legal compared to illegal nonwords). This absence of lexical effects is explained within the context of the race model by the suggestions that: (a) letter and letter cluster representations are faster to derive than lexical representations, and (b) fewer comparisons will be required to find a mismatch at these lower levels of representation for ‘different’ item pairs, since the comparison process is self-terminating and on average, only roughly half of the string will need to be processed before a mismatch is detected. For ‘different’ decisions, therefore, it is this reduction in the number of comparisons required at the lower levels of analyses that allows these levels to complete faster than the lexical level of analysis, and in-so-doing, prevents the expression of frequency and lexicality effects. The persistence of the orthographic effect for ‘different’ decisions suggests that the trade-off between time needed to extract a representation and the number of comparisons required to make a decision is fairly close and there is not such a great saving in comparison time at these levels.

The same-different matching task has been used in a variety of linguistic domains to tap into ‘automatic’ higher level processing without requiring overt decisions related to the factor under consideration, such as whether a string of letters forms a word, or whether words form a grammatical or meaningful sequence. For example, in addition to using this task to investigate lexical effects as Chambers and Forster did, Meyer and Schvaneveldt (1971) also used the task in their seminal work on privileged access for semantically associated words (see above). The same-different matching task has additionally been used to investigate the word superiority effect (see Henderson, 1980 for a brief review) and syntactic and semantic processes via sentence matching (e.g., Clahsen, Hong, Sonnenstuhl-Henning, 1995; Freedman & Forster, 1985; Murray, 1982).

Given the clear evidence that the same-different matching task engages lexical processing of two simultaneously presented words, it provides an ideal vehicle for investigating parallel lexical processing. The vertical alignment used in the classic studies does not, however, allow foveal and parafoveal effects to be tested simultaneously. To achieve this, it was necessary to align the word pairs horizontally rather than vertically. While higher level effects have never been sought using this alignment before, if the same 'race' principles apply, then the pattern of effects should mirror those obtained for vertically aligned word pairs, with effects of frequency, lexicality and orthography all being present for 'same' decisions at least.

If this pattern is shown under presentation durations that allow for the foveal inspection of each word in turn, it should then be possible to see whether the effects persist under shorter presentation durations where the foveal inspection of each word in turn should be prevented. If the same patterns of effects can be found with such short presentation conditions, then this would imply that the two words were being processed in an overlapping fashion. Varying word length should also allow an insight into whether ease of processing influences the likelihood that lexical effects will arise, with results from long words under short presentation durations providing the strongest test of whether the word pairs had been lexically processed in an overlapping fashion.

Participants in this study were therefore presented with spatially adjacent word pairs, which were either identical or differed by just one character. These pairs comprised either 4- or 6-letter words, of high or low frequency, or legal or illegal nonwords. In order to simulate the way in which words would normally appear in text, word pairs were always separated by a single character space.

Presentation times were varied to allow either the potential fixation of both words, with an eye movement (514ms), or to prevent one (216ms). The duration set for the short presentation condition was based upon the commonly-reported average fixation duration during reading falling within the range 225ms to 250ms (Rayner, 2009) and the eye-to-brain lag approximating 50ms (Foxe & Simpson, 2002). So even if the shortest average fixation duration is assumed, this, combined with the time require to propagate information to the visual cortex, should still be 275ms - 20% over the minimum duration set here.

For tasks similar to the one used here and where the fixated word remains visible, there is a tendency for even longer latencies. For example, Balota and Rayner (1983) found that the average saccade latency from a central to a peripheral word was 287ms, despite the central word being irrelevant to the task. But here, of course, the central word was task relevant; and therefore it appears reasonable to assume that even this underestimates the likely saccadic latency in the present experiment.

In order to ensure that participants were initially fixating the centre of the screen and not directing either overt or covert attention elsewhere, one word was always presented aligned with a fixation mark in the centre of the screen, while the peripheral word was presented to either its left or to its right with equal frequency. Such a procedure also has the advantage of preventing anticipatory saccades, which in turn should increase saccadic latency. Additionally, it allows assessment of the extent and direction of the perceptual span in this task. As will be recalled from Chapter 1, McConkie and Rayner (1976) showed that during natural reading in English, the perceptual span extends further to the right (15 characters) than to the left (3-4 characters). Consequently, we might expect to find stronger or more wide-ranging effects when the parafoveal word is presented on the right rather than when it appears on the left.

Any differences in matching time between legal and illegal nonwords would implicate sub-lexical orthographic processing having an effect; differences between legal nonwords and words with an entry in the lexicon would suggest that at least some aspects of lexical processing were involved; while differences in matching time due to word frequency, suggest that full lexical identification has taken place. Critically, in light of the current research question, if an effect of frequency is present under the short presentation durations, especially when the word pairs comprise two six-letter words, then this would imply that both words must have been processed, to a lexical level, in an overlapping fashion.

3.2. Method

3.2.1. Participants

Sixty-four native English speakers with normal or corrected to normal vision were tested. They received course credit or £5 payment for their participation.

3.2.2. Materials and Design

Four item conditions were used: high frequency words (e.g., ‘army’), low frequency words (e.g., ‘womb’), legal nonwords (e.g. ‘lumo’) and illegal nonwords (e.g., ‘qwdi’). Frequency was assessed using the Kuçera and Francis

(1967) norms. Words in the high frequency condition ranged from 100 to 1290 occurrences per million with a mean and standard deviation of 232.45 and 170.66, respectively; low frequency words ranged from 1 to 10 occurrences per million, with a mean and standard deviation of 5.60 and 3.33, respectively. There were 128 items of each type, comprising an equal number of four and six letter items.

Each item was tested in both the ‘same’ and ‘different’ conditions. For items requiring a ‘same’ decision, the item was repeated. For items requiring a ‘different’ decision, the item was matched with a comparison item differing by one character; the location of the differing character occurred with roughly equal frequency in each letter position. In the two real word conditions, to prevent discrimination based solely on word status, half of the differing items formed actual words while half did not. A full list of the items can be viewed in Appendix C.

Each item pair was displayed for either a short (216ms) or long (514ms) duration and the peripheral word appeared either to the left or right of the central word. This resulted in an 8 file counterbalanced design with each file containing an equal number of the 2 response types, 2 presentation durations and 2 presentation positions, all of which varied within items and an equal number of the 4 item types and 2 lengths, which both varied between items.

This resulted in each file containing 8 items in each of the 64 cells of the design.

3.2.3. Apparatus

Items were presented on a VDU screen using the DmDx program (Forster & Forster, 1997). DmDx utilises a High Performance Timer (accurate to within 1ms), which allows any trial with a display change error (i.e., a change that has missed its allotted refresh cycle) to be detected and flagged so that it can be disregarded during the latter stages of analysis. Furthermore, response time errors have a low probability of occurrence in DmDx and even when they do arise, they should not exceed 3ms.

All items were presented in white text on a black background using Courier New (size 10) font; letters were always in lower case. At the viewing distance of approximately 550mm there were approximately 2.25 characters per degree. Responses were recorded using left- and right-hand buttons on an attached button box

3.2.4. Procedure

Participants were told that they should initially fixate a cross, which would appear in the centre of the screen, and that this would then be followed by the simultaneous presentation of two horizontally aligned letter strings. They were told that the two letter strings might or might not be actual English words. It was made clear that word length, word position and presentation duration would vary from trial to trial, but that their task was simply to decide as rapidly and as accurately as possible whether the two strings were physically identical or not and to respond using the button box provided. They were informed that if the pair did differ, they would do so by just one character. A practice session with eight word pairs preceded the 512 experimental items.

Commencing each trial, a fixation cross appeared in the centre of the screen. After 1.7 seconds, the cross disappeared and the two letter strings were displayed. The middle of the central letter string was both horizontally and vertically aligned with the fixation cross; the peripheral word appeared either to its left or to its right separated by a single character space. In the ‘different’ condition, the central letter string was always the base word or nonword, while a differing letter string appeared in the periphery. To respond, participants pressed either a *same* (right) or *different* (left) button. After responding, they then pressed a request button to trigger the next trial.

3.3. Results and Discussion

As a result of the brief presentation and side-by-side display, error rates were understandably high, but well below chance (50%) in the vast majority of cases. However, in order to ensure that participants were treating the task seriously rather than just guessing, those with an overall error rate of 34% or higher were replaced. This was necessary on only three occasions. Reaction time cutoffs were set at 2.5 standard deviations from the mean reaction time, calculated individually for each participant. Any reaction time falling outside that range was replaced by the relevant cutoff value. Cutoffs were applied to only 1.8% of the reaction time data.

Response times and error rates were analysed for ‘same’ and ‘different’ responses separately. In each case, a 4x2x2x2 analysis of variance was carried out, treating subjects (F1) and items (F2) as random variables and including file as a between-groups dummy variable. The four factors were (a) item type (high and low frequency words; legal and illegal nonwords), (b) word length (4- vs. 6-letters), (c) presentation duration (short vs. long) and (d) presentation position (left vs. right).

An effect of item type was followed up with planned contrasts to determine whether it was driven by an effect of frequency (high vs. low frequency words), lexicality (low frequency words vs. legal nonwords), or orthography (legal vs. illegal nonwords), or a combination thereof. In the following analyses that include the four item types, where there was a violation of sphericity the Greenhouse-Geisser correction is reported.

3.3.1. ‘Different’ Responses

Mean response times and error rates for ‘different’ items are presented in Figure 3.1. An effect of item type was seen in both reaction times ($F1(1.64,91.54)=11.50$; $p<.001$; $F2(3,448)=7.38$; $p<.001$) and error rates

(F1(1.99,111.57)=93.68; p<.001; F2(3,448)=36.01; p<.001). It is apparent from the reaction time analysis that while there was little variation between the high and low frequency words and legal nonwords (all Fs<1), there was a clear orthographic effect in which legal nonwords were responded to faster than illegal nonwords (F1(1,56)=19.48; p<.001; F2(1,224)= 14.38; p<.001). This pattern of effects with 'different' responses is similar to those reported by Chambers and Forster (1975) who also obtained an orthographic but not a lexical effect when word pairs differed by a single character (Exp 2) and who also failed to uncover frequency effects with their 'different' items (Exp 1).

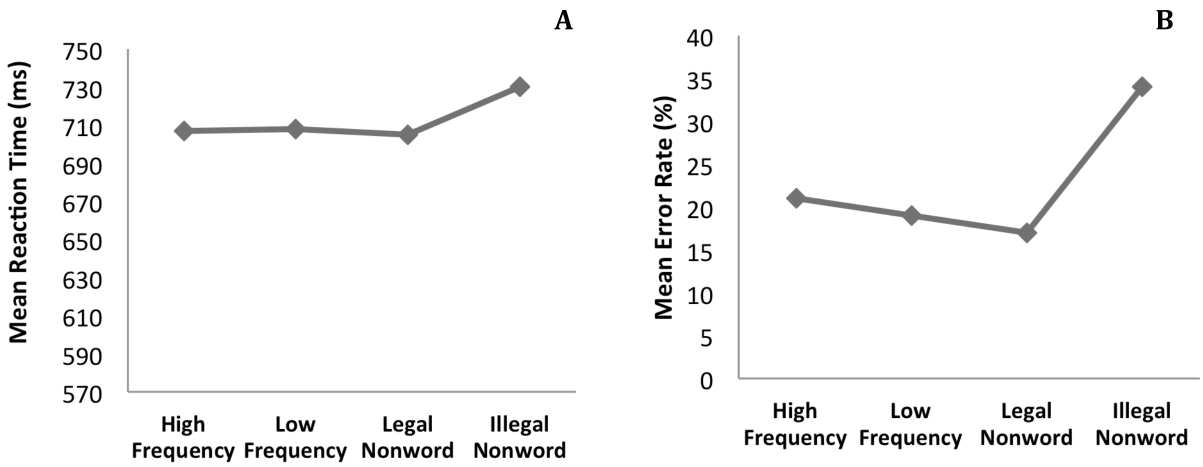


Figure 3.1. *Mean (A) Response Time (ms) and (B) Error Rate (%) for 'Different' Items as a Function of Item Type.*

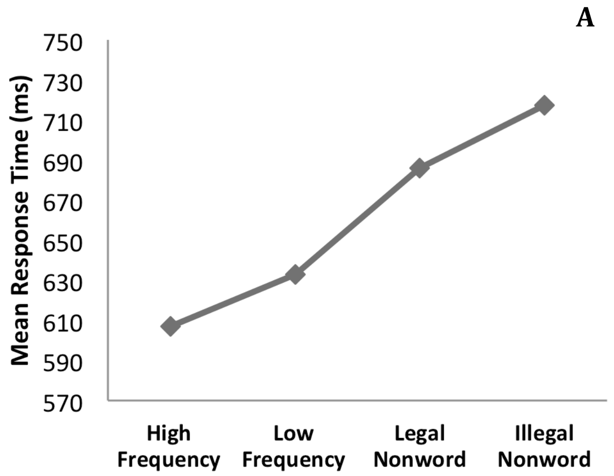
For error rates, there was an effect of frequency by-subjects (F1(1,56)=6.27; p<.05) but not by-items (F2(1,224)=1.75; p=.19), and an effect of lexicality by-subjects (F1(1,56)=6.97; p<.05) but again not by-items

(F2(1,224)=1.66; p=.20). This trend appears to have been driven by a response bias, in which participants were more likely to incorrectly respond ‘same’ the more familiar or ‘word-like’ the letter string was (see results for ’same‘ items below for the reverse trend). A clear effect of orthography was, however, present, with double the error rate for illegal letter strings compared to legal letter strings (F1(1,56)=200.00; p<.001; F2(1,224)=71.26; p<.001).

These results suggest that ‘different’ decisions were not based on comparisons made at a lexical level of analysis; rather sub-lexical ’letter clusters‘ appear to have been identified and compared, which resulted in fewer comparisons and therefore faster reaction times than when the comparison process could only be completed on a letter-by-letter self-terminating basis, as would be the case with the illegal letter strings. Such an interpretation of this orthographic effect is consistent with the increased error rates for conditions that would occasionally prevent the completion of multiple comparisons, such as when presentation durations were short (26% vs 20%; F1(1,56)=47.91; p<.001; F2(1,448)=83.82; p<.001) or when word length was long (28% vs 18%; F1(1,56)=180.37; p<.001; F2(1,448)=61.12; p<.001. Given that the exact processes underlying these ’different’ responses were not central to the current study, the remainder of this section will concentrate on ’same‘ responses, where lexical effects are normally found and are also apparent here.

3.3.2. 'Same' Responses

Effects of Item Type: As can be seen in Figure 3.2, both mean reaction times and error rates clearly show a main effect of item type ($F1(1.97,110.44)=209.91$; $p<.001$; $F2\ (3,448)=215.18$; $p<.001$; and $F1(1.83,102.42)=115.33$, $p<.001$; $F2(3,448)=136.19$; $p<.001$, respectively). Pairwise comparisons revealed that high frequency words were responded to faster ($F1(1,56)=76.57$; $p<.001$; $F2(1,224)=37.17$; $p<.001$) and received fewer errors ($F1(1,56)=44.85$; $p<.001$; $F2(1,224)=40.62$; $p<.001$) than low frequency words, which in turn received faster response times ($F1(1,56)=174$; $p<.001$; $F2(1,224)=109.45$; $p<.001$) and lower error rates ($F1(1,56)=175$; $<.001$; $F2(1,224)=79.22$; $p<.001$) than legal nonwords. Finally, legal nonwords were responded to faster ($F1(1,56)=48.30$; $p<.001$; $F2(1,224)=33.68$; $p<.001$) and were more error prone ($F1(1,56)=10.53$; $p<.01$; $F2(1,224)=10.27$; $p<.01$) than the illegal nonwords. Unlike the 'different' items, it is clear that effects of frequency, lexicality and orthography are apparent for items requiring a 'same' response'.



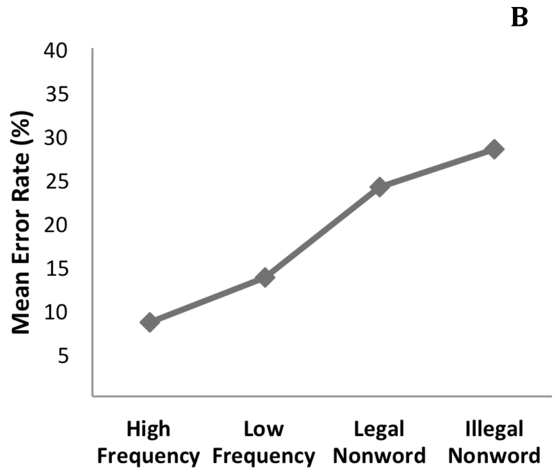


Figure 3.2. *Mean (A) Response Time (ms) and (B) Error Rate (%) for 'Same' Items as a Function of Item Type.*

Since the central word was identical for 'same' and 'different' items (in 'different' items, it was always the peripheral word that had been changed), the effects of frequency and lexicality seen here cannot be attributed to foveal processing alone, or the same pattern of results should have also arisen in the 'different' items. It is therefore apparent that characteristics of the parafoveal word must have contributed to this pattern of effects. We can determine whether this implicates parallel lexical processing by examining how item type interacted with presentation position, word length and presentation duration.

Effects of Position: Right-sided word pairs were responded to significantly faster (651ms vs. 671ms; $F1(1,56)=32.37$; $p<.001$; $F2(1,448)=56.82$; $p<.001$) and with fewer errors (17% vs. 20%; $F1(1,56)=15.39$;

p<.001; F2(1,448)=26.42; p<.001) than left-sided word pairs. As Figure 3.3 indicates, position and item type did not interact in the error rate analysis (F1(3,168)=1.70; p=.18; F2(3,448)=1.32; p=.27), but they did in the reaction time analysis (F1(2.66,148.77)=3.06; p<.05; F2(3,448)=3.75; p<.05). Follow-up analyses of the reaction time data revealed that this interaction was not driven by position interacting with either lexicality (both Fs<1) or orthography (F1(1,56)=1.59; p=.21; F2(1,224)=1.53; p=.22). Rather it appears to have been caused by an interaction between position and frequency (F1(1,56)=4.95; p<.05; F2(1,224)=3.90; p=.05), with a larger frequency effect when the peripheral word appeared on the right. However, the frequency effect remained reliable when the peripheral word was both on the left (F1(1,56)=22.24; p<.001; F2(1,224)=12.97; p<.001) and on the right (F1(1,56)=58.46; p<.001; F2(1,224)=36.98; p<.001).

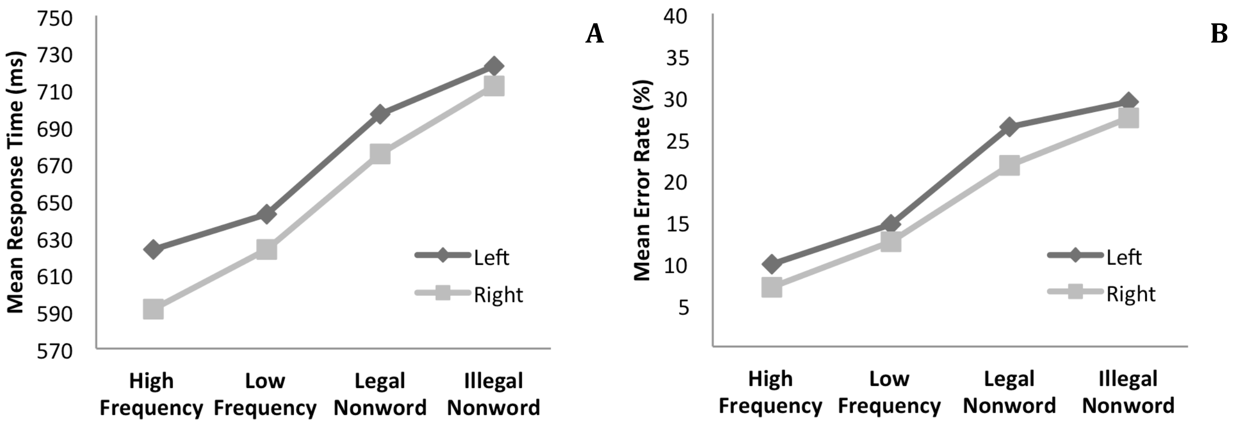


Figure 3.3. *Mean (A) Response Time (ms) and (B) Error Rate (%) for 'Same' Items for Each Presentation Position and as a Function of Item Type.*

This advantage in scanning to the right has been seen in numerous other studies (e.g., Rayner, 1978; Rayner et al, 1978; Balota & Rayner, 1983; Rayner et al, 1980; Simola et al, 2009) and appears to suggest that same-different matching engaged a similar asymmetrical distribution of attention as found in more natural reading (e.g., McConkie & Rayner, 1976).

Effects of Word Length: There was a main effect of word length, with 4-letter words responded to faster than 6-letter words (641ms vs 681ms: $F1(1,56)=160.07$; $P<.001$; $F2(1,448)=129.66$; $p<.001$) and generating fewer errors (16% vs 22%: $F1(1,56)=64.00$; $p<.001$; $F2(1,448)=57.43$; $p<.001$). As can be seen in Figure 3.4, word length did not affect the nature of the item type effect on reaction times ($F1(3,168)=1.20$; $p=.31$; $F2(3,448)=1.41$; $p=.24$); but there was a marginally significant interaction in the error rate by-subjects ($F1(3,168)=2.69$; $p=.05$; $F2(3,448)=1.54$; $p=.20$).

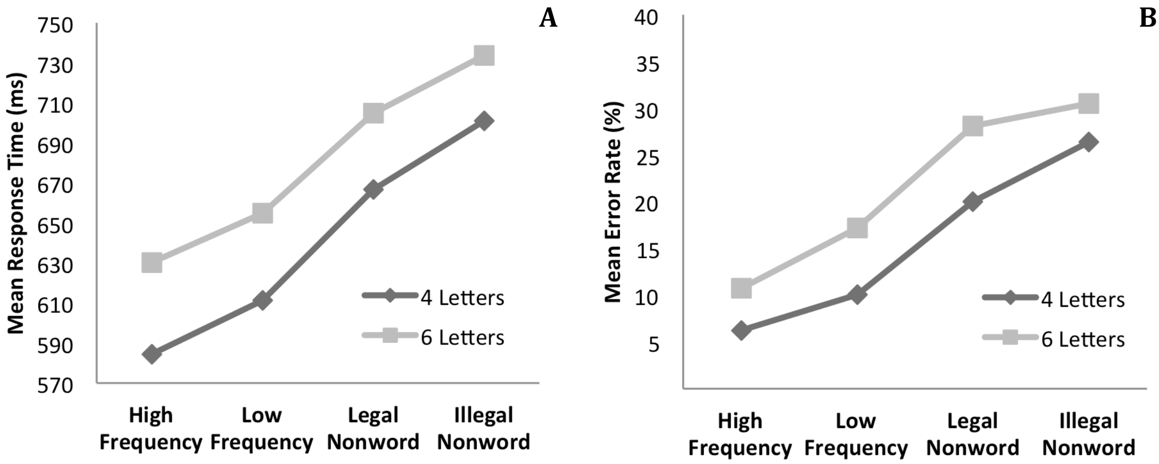


Figure 3.4. *Mean (A) Response Time (ms) and (B) Error Rate (%) for 'Same' Items for Each Word Length and as a Function of Item Type.*

As Figure 3.4B suggests, this interaction may reflect floor and ceiling effects on error rate, in which word length mildly interacted with frequency ($F1(1,56)=3.96$; $p=.05$; $F2(1,224)=2.67$; $p=.10$) and orthography ($F1(1,56)=5.02$; $p<.05$; $F2(1,224)=2.16$; $p=.14$), but not with lexicality ($Fs<1$). While the frequency effect persisted for both 4- ($F1(1,56)=32.08$; $p<.001$; $F2(1,112)=18.91$; $p<.001$) and 6-letter words ($F1(1,56)=26.13$; $p<.001$; $F2(1,112)=22.80$; $p<.001$), the effect size was slightly attenuated in the case of the shorter words, suggesting that the 6% error rate was the floor error rate for 4-letter high frequency words. The effective ceiling of just over 30% error rate might also explain why there was an orthography effect for 4-letter words ($F1(1,56)=20.73$; $p<.001$; $F2(1,112)=10.18$; $p<.01$) but not for 6-letter words ($F1(1,56)=1.70$; $p=.20$; $F2(1,112)=1.62$; $p=21$).

Floor and ceiling effects aside, it is clear that error rates follow the same general pattern as the reaction time data, and that even the longer 6-letter words show clear effects of frequency and lexicality, which suggests that even with 6-letter word pairs, sub-lexical and lexical information was being extracted and utilised from both words.

Effects of Duration: As can be seen in Figure 3.5, there was a speed accuracy trade-off in which short presentation durations resulted in faster responses than longer durations (650ms vs. 672ms; $F1(1,56)=18.94$; $p<.001$; $F2(1,448)=68.44$; $p<.001$) but they were more error prone (23% vs. 14%; $F1(1,56)=97.68$); $p<.001$; $F2(1,448)=209.74$; $p<.001$).

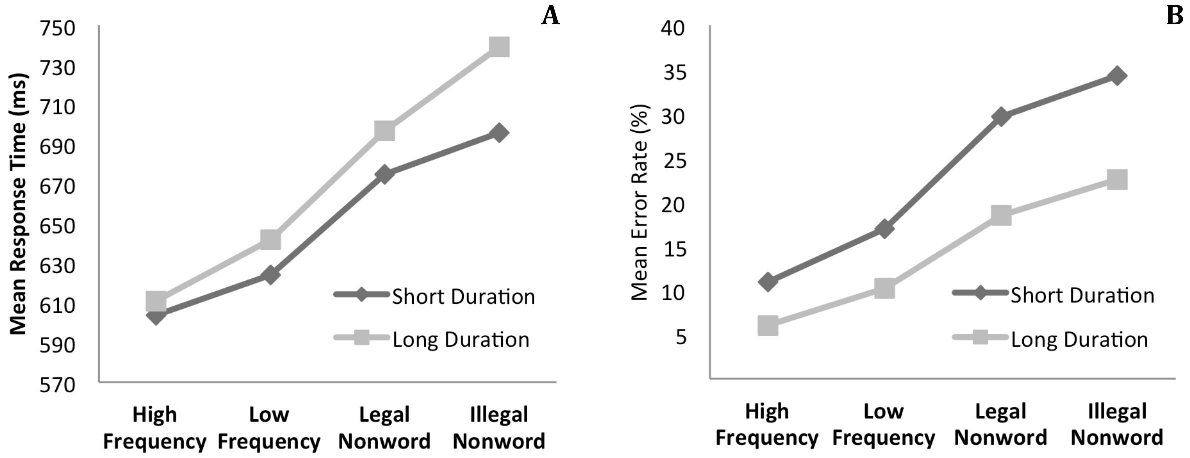


Figure 3.5. *Mean (A) Response Time (ms) and (B) Error Rate (%) for 'Same' Items for Each Presentation Duration and as a Function of Item Type.*

Duration interacted with item type in both the reaction time (F1(3,168)=9.11; p<.001; F2(3,448)=7.68; p<.001) and error rate analyses (F1(2,67,149.41)=7.20; p<.001; F2(3,448)=7.91; p<.001). Neither interaction appears to have been driven by duration interacting with word frequency (F1(1,56)=2.52; p=.12; F2(1,224)=2.42; p=.12; and F1(1,56)=1.88; p=.18; F2(1,224)=2.06; p=.15, respectively); with clear effects of frequency irrespective of duration in both the error rate (short duration: F1(1,56)=27.43; p<.001; F2(1,224)=27.98; p<.001; long duration: F1(1,56)=23.47; p=<.001; F2(1,224)=21.38; p<.001) and reaction time analyses (short duration: F1(1,56)=41.33; p<.001; F2(1,224)=36.53; p<.001; long duration: F1(1,56)=26.32; p=<.001; F2(1,224)=14.81; p<.001), suggesting that even with the shorter presentation condition, matching was performed at a lexical level.

As can be seen in Figure 3.5A, the interaction in the reaction time analysis was primarily driven by an attenuated effect of orthography when presentation duration was short ($F2(1,56)=7.23$; $p<.01$; $F2(1,224)=7.54$; $p<.01$). The effect of orthography, however, remained significant with short durations ($F1(1,56)=15.15$; $p<.001$; $F2(1,224)=8.54$; $p<.01$) as well as with longer ones ($F1(1,56)=40.62$; $p<.001$; $F2(1,224)=36.68$; $p<.001$). It therefore appears that when the most time-consuming comparison process is necessary to complete the task - that is, letter-by-letter comparisons – extended presentation time is utilised when it is available, while shorter presentations prevent this process from completing, encouraging a faster response.

The equivalent interaction was not, however, present for error rates (both $Fs<1$), with the interaction in the error rate analysis apparently being driven by an interaction between lexicality and presentation duration ($F1(1,56)=5.53$; $p<.05$; $F2(1,224)=6.40$; $p<.05$); strong effects of lexicality were however present when the presentation duration was short ($F1(1,56)=83.63$; $p<.001$; $F2(1,224)=60.37$; $p<.001$) as well as when it was long ($F1(1,56)=63.13$; $p<.001$; $F2(1,224)=42.60$; $p<.001$). Thus the patterns reported here suggest the ability to match at a lexical level appears to have provided an advantage when presentation duration was short.

Since the frequency, lexicality and orthographic effects remained significant for both presentation durations, it appears that even in those cases in which display duration precluded an effective second fixation, information from both words was still being utilised. Critically, the lexical effects appear to be undiminished at the shorter presentation duration.

In addition, the clear absence of a three-way interaction between item type, word length and presentation duration in both reaction time ($F1(3,168)=.60$; $p=.62$; $F2(3,448)=2.11$; $p=.10$) and error rate ($F1(3,168)=.30$; $p=.83$; $F2(3,448)=.35$; $p=.79$), suggests that the generality of these effects over different presentation durations holds for both 4- and 6-letter word pairs.

3.4. General Discussion

This study aimed to investigate whether parallel lexical processing is possible. In the same-different matching task employed here, it is clear that the comparisons were based on automatically derived higher-level lexical representations. Given that these effects were present with presentation durations designed to prevent an effective second fixation taking the eye to the parafoveal word, this implies that the two words were indeed being lexically processed in an overlapping fashion.

Taken as a whole, the results were relatively straightforward: Despite the very brief display and use of the side-by-side format, the patterns for the ‘same’ decisions replicate those found in classic same-different matching studies. Like earlier studies, there was clear evidence for a frequency effect (e.g., Chambers & Forster, 1975), a lexicality effect (e.g., Barron & Henderson, 1977; Barron & Pittenger, 1974; Barron, 1975; Chambers & Forster, 1975), and an effect of orthography (e.g., Chambers & Forster, 1975). As outlined above, Chambers and Forster (1975) proposed a 'race' model to account for this pattern of effects. This suggests that the comparison progresses at three levels of analysis simultaneously: the letter, letter-cluster and lexical levels, and whichever comparison completes the fastest wins the 'race' to determine the response. If a comparison can be made at a lexical level, this will likely complete the fastest because despite that representation taking the longest to derive, only one comparison will be required, and this appears to be reflected in the expressions of effects of frequency and lexicality in the present set of data. Equally, since fewer letter-cluster comparisons will be required than letter comparisons, this model can also explain the effects of orthography seen here.

In line with previous research, it appears that 'different' decisions were not usually made on the basis of higher level lexical representations; again, mirroring the findings of Chambers and Forster (1975), and it appears feasible to attribute this, as they did, to differences in the level of representation most likely to contribute to a 'no' decision most quickly. It will be recalled that the race model accounts for this pattern of effects via a trade-off between the time required to extract a representation at each level of analysis, and the number of comparisons required to find the difference. If a mismatch is present, fewer comparisons will be required at the lower levels of analysis, operating on a self-terminating basis. This, combined with the reduced time required to obtain the representations at these levels, prevents the lexical level from winning the ‘race’, resulting in an absence of lexical effects. The presence of an orthographic effect for ‘different’ decisions implies that this trade-off nonetheless still favours the letter cluster level, presumably since the time required to extract the letter clusters and make the comparisons is more time effective than when multiple letter level comparisons are required to find the mismatch.

The differences between ‘same’ and ‘different’ item types were anticipated; indeed, they nicely replicate those obtained by Forster and Chambers (1975), despite the difference in display format, and fit well within the conceptualisation of the race model. Since the ‘different’ item types do not exhibit any lexical effects, the remainder of this discussion will be devoted to the ‘same’ items, in which robust lexical effects were observed.

This study diverged from the classic same-different matching studies in that the words were horizontally and not vertically aligned. This adjustment allowed the two words to be presented in a format more akin to natural reading. The same patterns of ‘automatic’ higher-level effects were, however, obtained here, with clear effects of frequency, lexicality and orthography all surviving this change of format[19]. It was also evident that the horizontal alignment encouraged a similar asymmetrical span of attention to that found in natural reading (e.g., McConkie & Rayner, 1975). This replicates the findings from other isolated word or word pair tasks (e.g., Rayner, 1978; Rayner et al, 1978; McConkie & Rayner, 1983; Rayner et al, 1980; Simola et al, 2009) with a processing advantage associated with scanning parafoveal words to the right. Thus, it appears that many elements of reading have been automatically engaged with this task, making it an ideal vehicle for determining whether two words could in fact be lexically processed in parallel.

One potential concern with assuming overlapping lexical processing during a single fixation in this study is that eye position was not tracked. Therefore, it may have been the case that – even when presentation duration was just 216ms – participants extracted information from the foveal word and then made a quick saccade to the parafoveal one, processing each, foveally, in turn. While it is acknowledged that this must remain a possibility, previous research on saccadic latencies suggests this scenario is extremely unlikely. For example, Becker and Jürgens (1979) demonstrated that the average saccadic latency to a parafoveal target is typically between 175ms and 200ms. However, other studies - arguably more akin to this one - have reported average saccade latencies as low as 161ms (e.g. Rayner et al 1978). Therefore,

[19] Henderson (1974) also used horizontal alignment, but his items were restricted to orthographically illegal letter strings that were either meaningful or not meaningful (FBI vs IBF), for which he obtained a ‘word’ superiority effect for the former class of words.

adopting a more conservative line, if one adds 161ms to the 50ms eye-to-brain lag (Foxe and Simpson, 2002), information from a parafoveal word could be processed foveally within 211ms, which is roughly equivalent to the 216ms exposure time used in this study. However, neither Becker and Jürgens (1979) nor Rayner et al (1978) required participants to extract foveal information, and all the evidence suggests that the presence of a foveal word is likely to delay the initiation of a saccade. Indeed, Balota and Rayner (1983) found that when a word is placed in foveal view, saccadic latency to a parafoveal word is influenced by the lexical status of the foveal word, with longer latencies with real words (287ms) compared to nonwords (255ms). They surmise that "…although subjects were not required to read the foveal prime, its lexicality influenced their latencies to make a saccade" (pp. 729). And it should also be borne in mind, of course, that reading of the foveal stimulus was critical in this study, which is presumably likely to delay saccadic onset further. Consequently, while theoretically possible, it seems extremely unlikely that serial fixations were engaged with the shortest duration used here.

It might be suggested that an obvious next step would be to attempt to replicate these findings while tracking eye position. But while this would certainly help to address potential explanations in terms of serial processing associated with an overt eye movement, it would nonetheless fail to address another potential serial explanation: that a covert attention shift allowed the serial processing of the two words. It will be recalled from Chapter 1 that such covert shifts of attention are an essential feature of models like E-Z Reader (Reichle et al, 2009), and these might therefore appear to provide a plausible serial explanation for the present set of results; with information from the foveal word extracted prior to an attention shift to the parafoveal word. But while such an explanation is possible, it again seems rather unlikely.

In a review of the research into eye movement control during reading, Rayner (2009) reported that, on average, fixation durations range from 225ms to 250ms. Assuming that this average reflects the average time required to lexically process a word[20], then the central word in this experiment could feasibly have been lexically processed within 216ms. It has also been shown via gaze contingent reading experiments that reading can progress relatively smoothly when the foveal word disappears 50ms-60ms after the onset of a fixation on the word (e.g., Rayner, Inhoff, Morrison, Slowiaczak & Bertera, 1981), therefore it is possible that, following the shift of attention, enough information could have been extracted from the parafoveal word prior to it disappearing for that word's visual features to be stored and then subsequently lexically processed. Indeed, this seems to have been the case in the disappearing text experiments, where fixations falling on the blank space

[20] This is in fact the assumption of the most prominent models of eye movement control during reading (i.e., E-Z Reader and SWIFT). With saccadic programming running in parallel to lexical identification, *average* fixation duration should reflect the average time required to process a word, allowing some variability for preview benefit and spillover effects.

where the word had been continue to reflect the frequency of those words (e.g., Rayner, Liversedge, White & Vergillino- Perez, 2003).

There are, however, several reasons why this scenario seems unlikely in the present experiment. First, in the disappearing text experiments, the text disappeared following an overt eye movement, not a covert shift of attention. Acuity constraints will no doubt increase the time required to obtain visual information from a parafoveal word beyond the 50ms reported in the disappearing text studies. Second, This scenario assumes that attention shifts are instantaneous during reading – an assumption that is not uncontended (e.g., Inhoff, Radach & Eiter, 2006), Third, even allowing for such a scenario, the timeframe required to (a) lexically process the foveal word, and (b) extract visual information from the parafoveal one, should still exceed the time imposed by the short presentation durations in the current study. And, even if this process was occurring infrequently but still often enough to be driving the frequency effect, then one should expect to find a smaller frequency effect for the short compared to the long presentation durations, since the longer presentation did allow time for the sequential processing of two words on every trial. There was, however, little evidence for such an interaction. Thus, while an attention shift might have been responsible for these effects, the evidence would seem to suggest this is rather unlikely.

The only 'natural reading' study bearing some similarity to this was conducted by Inhoff, Starr and Schindler (2000), who included a word repetition condition (e.g., 'mother's mother'); they also had an associated condition (e.g., 'mother's father') and an unassociated condition (e.g., 'mother's garden'). These words were embedded within sentences with no contingent changes implemented. They found that gaze duration was shorter on the foveal word when the parafoveal word was related or associated than when it was unassociated. It might therefore be suggested that the word repetition employed in the current study helped facilitate the occurrence of parallel lexical processing as it appears to have done in the Inhoff et al study. Thus, there is some evidence that the results reported here loosely translate to a more natural reading task; although it is also acknowledged that the sort of word repetition employed here and in Inhoff et al's study is not typical of natural text, and it is not clear the extent to which possible parallel processing might depend on that. Therefore, while no assumptions can be made regarding the generalizability of these results, they nevertheless provide some evidence to suggest that multiple words can potentially be processed to a lexical level simultaneously.

Naturally, there is an argument that participants might employ task dependent strategies in studies such as these (e.g., Pollatsek et al, 2006; Rayner et al, 2003). A frequently referred to example of how non-reading tasks fail to capture the processes underlying natural reading is the finding that the frequency effect disappears during a search task (Rayner & Raney, 1996). But while the automaticity of reading is not apparent in every task where a word acts as the primary stimulus (e.g. visual search), this appears to be the exception rather than the rule. Same-different matching is clearly sensitive to higher level linguistic processes, not because of a task-dependent strategy, but simply because it turns out to be fastest to make the decision on the basis of (automatically generated) higher level representations. Certainly, for same-different matching tasks such as the one reported here, the patterns of results do closely mirror many of those from reading studies, suggesting that reading appears to be the default process across many tasks. All tasks involving reading carry a compromise in terms of validity. For example, participants may adjust their natural reading behaviour in eye tracking experiments involving sentence reading due to task instructions (e.g., being told to 'read for comprehension' and answer comprehension questions) and the environment these tasks necessitate (e.g., being monitored, biting on a bar, keeping head still, etc.). Thus, the utility of tasks such as the one reported here should not be dismissed simply because participants are not required to 'read sentences'.

It is acknowledged that the present experiment cannot adjudicate between serial and parallel models of eye movement control. These models were built to simulate the reading of sentences and not to make predictions about tasks such as this where reading happens to be automatically engaged. Indeed, that was not the purpose of this study. Its aim was to determine whether it is psychologically plausible to propose that two words could be lexically processed simultaneously. It did not aim to show that such a process is necessarily engaged during natural reading – just that it *could* be. At the very least, this study has demonstrated that two words can be lexically processed within the shortest fixation duration and most likely in an overlapping fashion. Certainly, both of those lexical representations were simultaneously available for comparison within that time scale. If the lexical processing of two adjacent words was completed in a serial fashion in that time during normal reading, then surely one would anticipate rather higher rates of word skipping.

CHAPTER 4

The Extent and Nature of Word N+2 Preview Benefit During Reading

4.1. Introduction

The two experiments reported in this chapter aimed to test the nature of word n+2 preview benefit during reading and the range over which it might be found. As will be recalled from Chapter 2, evidence pertaining to a word n+2 preview benefit has been mixed. While effects have been obtained, their expression has proved inconsistent. Amid several null reports (e.g., Angele Slattery, Yang, Kliegl, Rayner, 2008; Rayner Juhasz & Brown, 2007), some studies have reported word n+2 preview benefit, but - consistent with the premise that word n+1 identification precedes word n+2 pre-processing – only after word n+1 had been skipped (e.g., Angele & Rayner, 2011; Risse & Kliegl, 2011). Assuming that occasionally some of these intended skips fall short of their target – triggering a stay and process strategy (Drieghe, Rayner & Pollatsek, 2008) – then this also accounts for word n+2 preview benefit being expressed on word n+1 (e.g., Angele & Rayner, 2011; Kliegl et al, 2007; Risse & Kliegl, 2011). Therefore, as Schotter, Angele and Rayner (2012) note, these apparent parallel effects can be accommodated within a serial framework. Other effects do, however, seem less compatible. For example, Radach, Inhoff, Glover & Vorstius (2013) reported word n+2 preview effects on inspection of word n+2 when word n+1 had been fixated; moreover, these effects ‘spilt-over’ onto word n+3. Such positive results are, however, sparse, and replication appears essential.

4.2. EXPERIMENT 2

This experiment was designed to provide optimal conditions under which word n+2 preview effects might occur. As discussed in Chapter 2, previous failed attempts at uncovering word n+2 preview benefit within the English writing system might be linked to item selection issues. It has long been established that word length is positively correlated with fixation probability (Rayner, 1998), with a significant increase in fixation probability from 3- to 4-letter words (32% vs. 48%; Rayner & McConkie, 1976). Thus it is likely that the inclusion of 3-letter words in the position of word n (as was the case in: Angele et al, 2008; Angele & Rayner, 2011 and Rayner et al, 2007) might increase the probability that the fixation before the change actually fell on word n-1, resulting in the study actually testing a preview benefit (or lack thereof) related to word n+3 rather than word n+2.

Conversely, the inclusion of relatively long words in the positions of word n (Angele et al, 2008; Angele & Rayner, 2011), n+1 (Angele et al, 2008) and n+2 (Angele et al, 2008; Angele & Rayner, 2011; Rayner et al, 2007) might have prevented attention, on occasions, from reaching word n+2. As will be recalled from Chapter 1, the limit of the effective span of apprehension typically falls around the 14th or 15th character space to the right of fixation (McConkie & Rayner, 1976), thus two or more relatively long words within the critical word sequence might drive word n+2 outside this window.

This conjecture is not new; Angele et al (2008) acknowledge that their use of longer words in the position of word n+1 (range 4- to 10-letters) may have contributed to their null result. While this was remedied in Angele and Rayner's (2011) study, in which word n+1 always contained just 3-letters, across experiments 1 and 2, words n and n+2 included words of up to 10- and 14-letters, respectively, which again may have resulted in word n+2 often falling outwith the effective span of apprehension. Rayner et al (2007) also acknowledged that the null results obtained in their first experiment might be attributed to critical word length. This was addressed in their second experiment. However, these words were typically low frequency, which may have unduly constrained attentional resources and again may have biased the results towards the null hypothesis. Indeed, there is evidence that attentional resources can be modulated by ease of processing, with low frequency foveal words apparently reducing word n+1 preview benefit compared to high frequency foveal words (Henderson & Ferreira, 1990; White, Rayner & Liversedge, 2005). Overall, it appears therefore that these studies do not provide optimal conditions under which extended parafoveal processing might be expected on the basis of previous research.

Radach et al (2007) proposed a set of criteria under which word n+2 preview benefit might optimally occur. First, they emphasized the importance of keeping all critical words high in frequency so as not to constrain attentional resources. Additionally, critical word lengths should be controlled such that: word n should be of a medium length in order to attract one fixation; word n+1 should be short to promote prompt visual and linguistic processing; and finally, word n+2 should again be of medium length in order to attract a fixation and to allow, at the very least, the beginning of the word to fall within the effective span of apprehension. While item sets used in German studies have varied, in line with Radach et al's proposal, positive evidence for word n+2 preview effects have typically arisen from German studies that (a) refrain from using 3-letter words in the position of word n, and (b) keep word n+1 to just 3-letters in length (e.g., Kliegl, Risse & Laubrock 2007; Radach, Glover & Vortsius, 2007; Radach et al, 2013; Risse & Kliegl, 2011).

This study therefore set out to determine whether word n+2 preview effects can also obtained using the English writing system when 'optimal materials' are employed. Angele and Rayner (2011) provide some evidence to suggest that this is possible, but only in those cases when word n+1 had been skipped, which allows the possibility that the effect may have arisen via a double-attention shift and would therefore be compatible with the E-Z Reader model. But, since it has been shown that word n+1 fixation is not always a necessary requirement for word n+2 preview effects when tighter control over the materials is exercised (Radach et al, 2007; Radach et al, 2013), it is possible that Angele and Rayner's effects were driven by a subset of the items and that the conditions required for word n+2 preview benefit following a fixation of word n+1 were absent from their item set (e.g., word n+2 was often too long to fall within the effective span of apprehension).

Based on these observations, all critical words in the present study were of a high frequency so that they were easy to process and unlikely to constrain attentional resources. Words n and n+2 were also of a medium length (6-letters), which should be long enough to attract a fixation[21] but also not so long that word n+2 would fall outwith the effective span of apprehension while fixating word n. Finally, word n+1 was always 3-letters in length.

One potential problem associated with including such short easy to process words in the position of word n+1 is that they might, within the architecture of the E-Z Reader model, permit a double attention shift to occur, with attention shifting serially from words n, to n+1 and onto n+2, all before a saccade is executed terminating the fixation on word n. Thus, despite providing optimal conditions for parallel lexical processing, such a choice of materials fails to provide a strong test in which the SWIFT and E-Z Reader models can easily be differentiated.

[21] According to research conducted by Rayner and McConkie (1976), 6-letter words are fixated on over 70% of occasions.

To remedy this, four preview conditions were employed such that prior to passing an invisible boundary located immediately after word n either: both words n+1 and n+2 received an identical preview, only word n+1 received an invalid preview (testing word n+1 preview benefit), only word n+2 received an invalid preview (testing word n+2 preview benefit, but with the potential for a double attention shift), or both words received an invalid preview. This latter condition is important since, according to the E-Z Reader model, a double attention shift can only arise if word n+1 can be identified. With an invalid preview of word n+1, attention cannot shift legitimately onto word n+2. Therefore, according to the E-Z Reader model, if word n+1 receives an invalid preview, the word n+2 preview manipulation should not influence the eye movement record.

Finally, similarly to Kliegl et al (2007) and Angele and Rayner (2011), word n+1 was either a determiner or an alternative high frequency 3-letter word. This manipulation was motivated by Radach's (1996) word grouping hypothesis, which proposes that determiner-noun pairs may be processed as one perceptual unit, in a manner distinct from other word pairs, while the alternative high frequency 3-letter word-noun pairs should be processed as two separate perceptual units. This theory was based upon Radach's observations that landing site distributions for determiner-noun pairs appear to share a single distribution similar to that typically observed on a single word, while the alternative word-noun pairs appear to show two separate distributions – one for each word (c.f., Drieghe et al, 2008, but even in that study similar distributional properties can be observed). If determiner-noun pairs are indeed processed as one perceptual unit, one might predict stronger word n+2 preview effects for these compared to alternative 3-letter high frequency word pairings. Past efforts to investigate the word grouping hypothesis within a word n+2 preview experiment have, however, proved inconsistent. Angele and Rayner (2011; Exp1) failed to observe any evidence that determiners and alternative high frequency words were grouped differently with their subsequent nouns; but, as discussed above, word n+2 preview effects might have been constrained in that experiment by item selection effects. Kliegl et al (2007) did, however, obtain results indicating that word n+1 type modulated a word n+2 parafoveal–on-foveal effect, with longer durations on word n when the upcoming word n+2 had received an invalid rather than identical preview, but only in those cases where word n+1 had been a content and not a function word. This pattern falls in the opposite direction to the predictions of the word grouping hypothesis, where facilitated word n+2 processing should have been encouraged for determiner-noun pairs and not alternative word-noun pairs. Given these inconsistencies, the present study sought to further investigate the word grouping hypothesis.

To summarise, the primary motivation for Experiment 2 was to investigate word n+2 preview effects within the English writing system under optimal conditions related to critical word length and frequency. If, however, these effects remain elusive, or indeed if they only arise when word n+1 receives an identical (i.e., not invalid) preview, then this would be consistent with the deployment of a serial attention allocation process during reading. Further, any word n+2 parafoveal-on-foveal effect expressed on word n would strongly suggest that attention is being distributed in parallel fashion. This is because the mechanism the E-Z Reader model utilises to account for parafoveal-on-foveal effects during reading is a saccadic undershoot of a target word followed by a stay and process response (see Chapter 2 for more details); thus, to account for a word n+2 parafoveal-on-foveal effect on word n within E-Z Reader's architecture, one must assume that targeting errors of 5 or more characters are prevalent in the eye movement record, which is unlikely to the case. Finally, based on the word grouping hypothesis, word n+2 pre-processing should be facilitated for determiner-noun pairs compared with alternative high frequency word-noun word pairs.

In addition to investigating a range of word-based measures on words n, n+1 and n+2, fixation time and saccade targeting measures were also analysed for a 3-word spillover region. Analyses on this region were included to determine whether, like Radach et al (2013), any delayed effect of the preview manipulations was present further 'downstream'. To avoid the potential confound of sentence wrap-up effects (Aaronson & Scarborough, 1976; Just & Carpenter, 1980; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989), this region was never placed at the end of a sentence.

4.2.1. Method

4.2.1.1. Participants

Sixty-four native English speakers with normal or corrected-to-normal vision and with no known reading difficulties took part in the experiment. Each received course credit for their participation.

4.2.1.2. Materials and Design

Forty-eight experimental items were constructed. Each sentence frame contained an initial noun phrase followed by a 6-letter verb – designated as word n. This was followed by either a determiner or 3-letter high frequency word (n+1) and then by a 6-letter noun (n+2)[22]. Words n, n+1 and n+2 lengths were chosen to encourage attention to stretch to word n+2 while fixating word n. For this reason these words were also high frequency (word n: M=135, SD=98; word n+1 determiner: M=69,971, SD=0; word n+1: alternative: M=3040, SD=2266; word n+2: M=134, SD=90; all values per million estimated using the Kuçera & Francis, 1967, norms). Sentences ranged from 60 to 85 characters in length, including spaces.

There were 8 versions of each of these experimental items presented to participants. Words n+1 and n+2 were initially presented in parafoveal vision in either their correct form or as a nonword of equivalent length – an 'invalid'

[22] After testing was complete, it was noted that 1 verb and 1 noun in fact contained 7-letters. Since these words occur in all conditions, they were retained in the analyses.

preview. Nonwords typically formed orthographically illegal letter strings that matched their respective target word's word envelope. Thus there were 4 experimental preview conditions in total with one, the other, neither or both of words n+1 and n+2 initially receiving an invalid preview.

The boundary used to trigger the contingent change always occurred immediately following the verb (word n). Crossing this boundary triggered a display change to occur, presenting words n+1 and n+2 in their 'correct' form. As can be seen in Figure 4.1 below, this resulted in a 2 (word n+1 type) x2 (word n+1 preview) x2 (word n+2 preview) design with a total of 8 conditions. The full set of experimental items can be seen in Appendix G.

Determiner n n+1 n+2

1a) The book editor forced|(the writer) <u>the writer</u> to rename the controversial book.

1b) The book editor forced|(fbo writer) <u>the writer</u> to rename the controversial book.

1c) The book editor forced|(the mochsv) <u>the writer</u> to rename the controversial book.

1d) The book editor forced|(fbo mochsv) <u>the writer</u> to rename the controversial book.

HF Word n n+1 n+2

1d) The book editor forced|(his writer) <u>his writer</u> to rename the controversial book.

1e) The book editor forced|(kwe writer) <u>his writer</u> to rename the controversial book.

1f) The book editor forced|(his mochsv) <u>his writer</u> to rename the controversial book.

1g) The book editor forced|(kwe mochsv) <u>his writer</u> to rename the controversial book.

Figure 4.1. *Example item showing each of the 4 parafoveal preview conditions for each of the two n+1 word types (1a-1d determiners; 1e-1g: High frequency words). Parafoveal previews are presented in parentheses, while target words n+1 and n+2 are underlined. The boundary location is denoted by the symbol: "|".*

Eight counterbalanced item files were constructed. Each participant experienced all preview conditions across an equal number of items, but saw only one version of each item. The particular allocations of items to files and participants to files were treated as between-groups dummy variables in the following analyses.

To ensure normal reading for comprehension, 10% of the experimental items were followed by a comprehension question. In addition, a further 56 similar items with 12 comprehension questions were constructed as filler items[23]. Eight separate practice items preceded the experimental items; half of these were accompanied by a comprehension question.

4.2.1.3. Apparatus

Participant's head-movements were constrained by use of a dental wax bite bar and chin rest and their eye-movements recorded using a Dr. Bouis pupil-centered computation Oculomotor interfaced to a 12-bit A-D device sampling X and Y position every 2ms. Viewing was binocular but only the movement of the right eye was monitored. Sentences were displayed in monopitch white text on a single line of a black screen on a CRT display refreshed at 100Hz, at a viewing distance of around 500mm. At this distance, one character subtended approximately 0.3 degrees of a visual angle. Participants answered comprehension questions by pressing either a right- or left-hand button (for 'yes' or 'no', respectively) on an attached button-box.

The contingent change of the target words was achieved by writing directly to the video memory of the graphics control card and was not dependent on the refresh cycle of the display (see Kennedy, Pynte & Ducrot, 2002, for details). This procedure ensured that the display change occurred

[23] Forty-eight of these formed the experimental items for the second experiment reported in this chapter.

within 13ms and the preview was displayed only while the eyes were to the left of the invisible boundary. To ensure strict comparability between display conditions, the contingent change was also employed in the identical preview condition (with the target letters replaced by themselves).

4.2.1.4. Procedure

On arrival, participants were told that their eye movements would be monitored while they read a set of sentences and answered associated comprehension questions. They were asked to read the sentences normally and then answer the questions as accurately as possible. Both verbal and written instructions were provided. The written instructions can be seen in Appendix E.

Calibration involved the fixation of five points distributed evenly across the horizontal axis of the screen at the point where the experimental sentences were to be displayed. Initial setup and calibration took approximately five minutes and to ensure accuracy, a brief re-calibration was repeated after every four sentences during the experiment.

Each trial began with the display of a fixation marker in the form of a small cross (+) located 3 character spaces to the left of where each sentence would begin. When the computer detected a stable fixation on the marker for at least 100ms, the marker disappeared and the sentence was displayed in a monopitch font on a single line of the CRT display. After reading each sentence, participants pressed a right-hand button. This resulted in the experimental sentence being replaced with either a comprehension question or a sequence of dashes. If dashes appeared, participants pressed the right-hand button again, or if a question appeared, they pressed either the right- or left-hand button to respond “yes” or “no” respectively; in both cases, the response triggered the beginning of the next trial. The entire experiment lasted approximately 60 minutes and participants were given a break whenever they desired.

4.2.2. Results and Discussion

As shown in the example below, for purposes of analysis, four zones were defined in each of the experimental items: one corresponding to each of the words n, n+1 and n+2 (zones 1, 2, 3, respectively), and a 3-word ‘spillover’ region (zone 4). Fixations falling on the space preceding each of these were also considered to have fallen into the relevant region.

```
Zones:               1      2     3          4

The book editor| forced| the| writer| to rename the| controversial
book.
```

A number of fixation time measures were computed for each of these zones. These included:

First and Last Fixation Durations: The duration of the first and last fixations falling in each zone during first-pass reading.

Single Fixation Duration: The duration of the first fixation falling in a zone providing that zone only received a single fixation during first-pass reading.

While first, last and single fixation durations constitute overlapping sets, they potentially tap into differing processes and are therefore reported independently.

***Gaze Duration*:** The summed duration of all fixations within a zone on the first-pass before the eye exited it either to the left or to the right. Since it is generally assumed that gaze duration reflects the time taken to identify an 'object', a skipped region will contribute a value of 0ms to the calculation of average gaze duration. As discussed in Chapters 1 and 2, there is now good evidence to suggest that gaze duration on a word does not solely reflect the time required to process that word. So-called 'spillover effects' and 'preview benefit' both suggest a somewhat looser coupling between fixation location and the locus of attention than Just and Carpenter (1980) envisaged. Given this, it seems reasonable to assume that if a word has not been fixated during the first pass, it was, at least partially, identified during the preceding fixation and therefore its processing time will be absorbed into the processing time of the preceding word. Additionally, of course, the processing of the region might also be continued during fixations that follow it. It is argued that if zeroes are excluded from gaze duration on the skipped word, then the processing time for that word will be counted twice: once in the preceding fixation, and once again – by means of an averaged replacement of the zero – on the word itself (see Murray, 2000 for further discussion). For this reason, the gaze durations reported here include zeroes. However, any effects pertinent to the research question that are clearly the result of word skipping will be highlighted and additional analyses will be reported in which the calculation of gaze duration excludes zeroes.

Go-Past Reading Time**:** The summed duration of all fixations from initial inspection of the zone on the first pass until the eye exited it to the right. Again, for the reasons outlined above, zeroes are included when calculating mean go-past time (but, like gaze, results from the alternative calculation will also be reported when required). It should be noted, however, that it is only with the inclusion of zeros that the summed go-past time for two adjoining regions will equal the go-past time calculated for a single region that encompasses both of these.

First-Pass Re-Reading Time**:** The summed duration of all fixations in regressions to prior regions or in re-reading of the region before the eye exited it to the right.

In addition, two saccadic measures will be reported for most regions:

Skipping Probability: The probability that a zone was skipped during first-pass reading (restricted to words n, n+1 and n+2).

First Landing Position: The position (in character spaces) of the first fixation falling within each zone during first-pass reading.

Analyses were performed treating both Participants (F1) and Items (F2) as random factors. These were repeated measures 2 (determiner vs. alternative high frequency word) x 2 (word n+1: identical vs. invalid) x 2 (word n+2: identical vs. invalid) analyses of variance (ANOVAs) conducted for each of the above measures for zones 1 to 4. In all analyses, file was treated as a between-groups dummy factor.

Participants achieved an overall accuracy on the questions of over 80%, suggesting they had read the sentences carefully.

4.2.2.1. Word N

Effects of Word N+1 Type: A significant effect of word n+1 type was apparent in first-pass re-reading time, with a 12ms increase when the word to the right was a determiner rather than an alternative high frequency word (35ms vs. 23ms: $F1(1,56)=4.25$; $p<.05$; $F2(1,40)=10.80$; $p<.01$). However, it can be seen in Table 4.1, that aside from a similar but numerically small, non-significant trend in last fixation duration (243ms vs. 239ms: $F1(1,56)=2.38$; $p=.12$; $F2(1,40)=2.08$; $p=.15$), there were no other main effects of word n+1 type in any other measure (all $Fs<1$). Therefore, similar to the findings of Kliegl et al (2007), readers did show sensitivity to the class of the upcoming word, although the response clearly differed between the two experiments. While

Kliegl et al reported an increase in gaze duration when the parafoveal word was a content word; the alternative high frequency word in the present experiment appears to have inhibited regressions. Inconsistencies in the expression of parafoveal-on-foveal effects, such as this, are not uncommon in the literature and potential explanations for this variation will be returned to in the Discussion. For now, it is important to highlight that since this effect was reflected in re-reading time alone it cannot be accounted for by a skipping cost or benefit associated with the upcoming word (e.g., Kliegl & Engbert, 2005). A mislocated fixations explanation (e.g., Drieghe et al, 2008) also fails to provide a plausible explanation since this would require an undershoot of word n+1 followed by a stay and process response – not a regression, as seen here.

Table 4.1. *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for Word N.*

	Determiner				High Frequency Word			
	N+1 Identical		N+1 Invalid		N+1 Identical		N+1 Invalid	
	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid
First Fix	244	242	245	244	243	243	241	243
Last Fix	243	242	239	245	235	242	236	239
Single Fix	250	247	246	250	250	247	245	249
Gaze	240	235	246	249	256	232	253	249
Go-Past	269	274	280	288	277	249	287	270
Re-Reading	29	39	34	39	21	17	34	21
Skip Prob	13	14	15	14	11	17	11	12
Landing	2.82	2.93	2.84	2.93	2.89	2.78	2.70	2.85

The E-Z Reader model could potentially be adapted to account for this result if it was assumed that a fast acting integration failure caused the eyes to regress back further than the word where the comprehension difficulty was first detected. It could be the case that word n+1 was fully identified while fixating word n, followed by a fast-acting integration failure of that word, triggering a regression out of the verb. The higher level processing module has only recently been implemented within the E-Z Reader model and the architects acknowledge that such regressions are currently underspecified in it;

however, it is conceivable that such an adaptation could be incorporated in a future version of the model. From a theoretical perspective, however, it seems counterintuitive to suggest that a determiner should be more difficult to integrate than an alternative high frequency word. A more parsimonious explanation appears to be that word n+1 was identified while fixating word n, which in turn influenced the reader's reading strategy.

Effects of Word N+1 Preview: The presence of an invalid preview to the right of fixation tended to increase both gaze duration (241ms vs 250ms: $F1(1,56)=3.59$; $p=.06$; $F2(1,40)=2.97$; $p=.09$) and go-past time (267ms vs 282ms: $F1(1.56)=7.05$; $p<.05$; $F2(1.40)=3.60$; $p=.06$) compared to when word n+1 was parafoveally available. These trends mirror those found in other studies (e.g., Angele & Rayner, 2011; Inhoff, Starr & Schindler, 2000; Starr & Inhoff, 2004) and achieved significance when zeroes were treated as missing data in the calculation of go-past time (303ms vs 319ms: $F1(1,56)=7.80$; $p<.01$; $F2(1,40)=4.06$; $p<.05$). Since these effects were not significant in any other measure (first-pass re-reading time: both $Fs<1.5$; all other measures: $Fs<1$), it seems that they were driven – at least predominantly - by an increased tendency to re-fixate word n when orthographic illegality was detected to the immediate right. These results add to the growing body of evidence to suggest that the orthographic properties of a parafoveal word can influence foveal inspection times. While such an effect was once considered problematic for serial models of eye movement control, it has recently been suggested that

that the E-Z Reader model can account for such effects via the low level attentional scan. The plausibility of such an auxiliary assumption is something that will be returned to in the Discussion.

Effects of Word N+2 Preview: There were no reliable main effects of word n+2 preview in any of the durational measures (first and single fixation durations, $Fs<1$; last fixation duration: $F1(1,56)=1.26$; $p=.27$; $F2(1,40)=2.21$, $p=.14$; gaze duration: $F1(1,56)=2.43$; $p=.12$; $F2(1,40)=1.30$, $p=.26$; go-past time: $Fs<1.5$; and first pass re-reading time: $Fs<1$), nor in first landing position (both $Fs<1$). A trend did emerge in skipping probability ($F1(1,56)=3.19$; $p=.08$; $F2(1,40)=3.19$; $p=.08$), with 2% more skips of word n if word n+2 was invalid in the parafovea (12% vs. 14%). However, as can be seen from Figure 4.2.A, this main effect appears to be modified by an interaction related to whether or not word n+1 was also previewed in its correct form ($F1(1,56)=2.81$; $p=.10$; $F2(1,40)=3.44$; $P=.07$). It appears that skipping of word n only increases in response to an invalid preview of word n+2 when word n+1 had been parafoveally available ($F1(1,56)=5.24$; $p<.05$; $F2(1,40)=7.72$; $p<.01$). Unsurprisingly, when N+1 was parafoveally invalid, there was no effect (both $Fs<1$). There was no evidence for any difference in this pattern across the two n+1 word types (determiner vs HF) with no three-way interaction (both $Fs<1$).

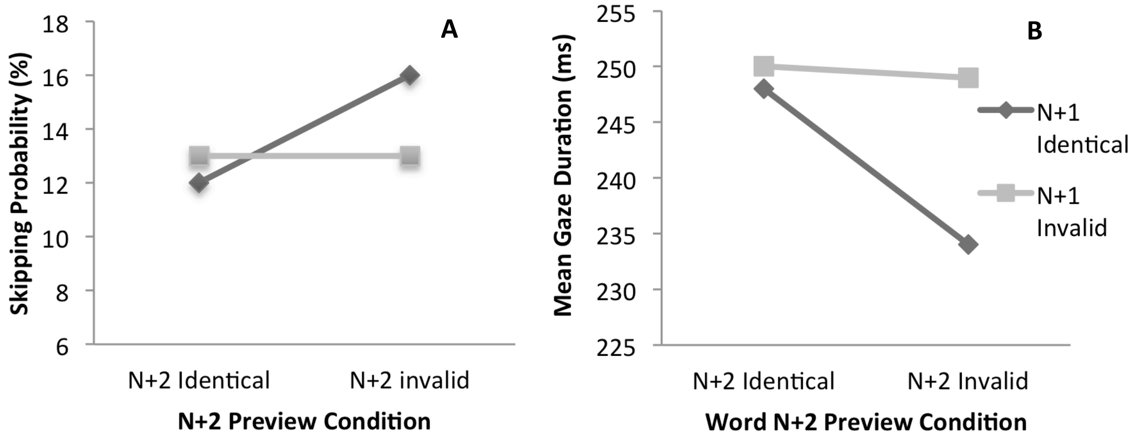


Figure 4.2. *Mean (A) Skipping Probability (%) and (B) Gaze Duration (ms) on Word N for the Two Word N+1 Preview Conditions as a Function of Word N+2 Preview.*

A similar, though nonsignificant, pattern can also be seen in gaze duration in Figure 4.2.B, (F1(1,56)=2.56; p=.11; F2(1,40)=1.84; p=.18), with a trend towards a reduction in gaze duration when word n+2 was invalid and word n+1 was available (F1(1,56)=4.12; p<.05; F2(1,40)=3.37; p=.07), but no difference when it was not (both Fs<1). This pattern appears to be due to skipping, since a skipped word is treated as having a gaze duration of zero. This conclusion is further verified by the fact that any evidence for an interaction between the two (n+1 and n+2) preview manipulations disappears when skips are treated as missing data in gaze (both Fs<1).

Since the decision to skip word n must occur while the reader is fixating word n-1, this potentially reflects a word n+3 parafoveal-on-foveal effect, to which the response appears to be for the eyes to move forward more than they otherwise would. Such ‘attraction’ effects are not new (e.g., Angele &

Rayner, 2011; Hyönä & Bertram, 2004) and would appear to present some difficulty for serial models, such as E-Z Reader.

It is however, conceivable that the skipping effect (and the corresponding effect in gaze duration) could be accounted for within the architecture of the E-Z Reader model if one assumes a triple attention shift. That is, word n was identified while fixating word n-1, allowing attention to shift onto word n+1 which – because it was parafoveally available - could in turn have been identified allowing attention to proceed once again to word n+2. This would necessitate that the reader, within the time required to program, cancel and reprogram a saccade out of word n-1 had completed the lexical identification of both word n and word n+1 and had then directed their attention to word n+2. Needless to say, such an account would necessitate the execution of multiple serial operations within an extremely tight time window. Whether or not this seems viable will be considered further in the Discussion.

An alternative appears to be to abandon the above line of reasoning and to rely instead on the low level attentional scan postulated by proponents of the E-Z Reader to account for the present pattern of effects. At present, however, there is no way to use the model to simulate these effects, so it is unclear whether they could be accounted for this way. Doing so would, however, probably involve postulating a rather broad effective range for the low-level scan. It should be noted that both potential explanations for these results appear to require a much looser coupling between fixation location and attention than Just and Carpenter (1980) envisaged; a concept from which the proponents of E-Z Reader originally drew inspiration.

Finally, as can be seen from Figure 4.3, although there was no significant main effect of word n+2 preview in go-past or re-reading time, these measures did show an interaction between the effect of n+2 preview and word n+1 type. This was significant in go-past time ($F1(1,56)=5.01$; $p<.05$; $F2(1,40)=7.54$; $p<.01$) and apparently driven, in part, by a trend towards an equivalent interaction in re-reading time ($F1(1,56)=2.53$; $p=.11$; $F2(1,40)=3.15$; $p=.08$). Pairwise comparisons revealed that when word n+1 was a determiner, there was no effect of N+2 preview (all $Fs<1$); but when word n+1 was an alternative high frequency word, the 18ms decrease with an invalid n+2 preview was significant in go-past time ($F1(1,56)=7.74$; $p<.01$; $F2(1,40)=5.73$; $p<.05$), with a similar 9ms tendency emerging in first-pass re-reading time ($F1(1,56)=2.39$; $p=.12$; $F2(1,40)=3.01$; $p=.08$).

Similarly to Kliegl et al (2007), readers did show sensitivity to the status of word n+2 preview when word n+1 was an alternative high frequency word, but again the expression of this effect differed between the two studies. While Kliegl et al reported increased inspection time on word n when word n+2 received an invalid preview, the present study found the opposite pattern. Potential explanations for this difference will be considered in the Discussion, but there is no immediately obvious explanation for such patterns of effects within the architecture of the E-Z Reader model, especially with the effects modulated by word n+1 type. These results, again, appear to implicate the sort of attraction hypothesis suggested by Kennedy (1998) and also reported by Hyönä and Bertram (2004), where an upcoming irregularity attracts attention to the region of difficulty.

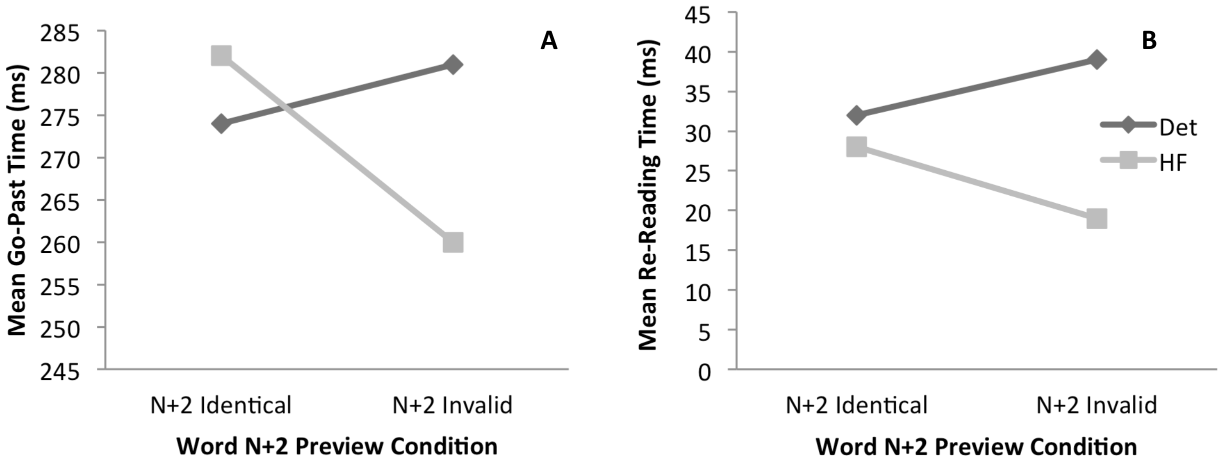


Figure 4.3. *Mean (A) Go-Past Time and (B) First-Pass Re-Reading Time (ms) on Word N for Determiners and High Frequency Words as a Function of Word N+2 Preview.*

There were no other significant interactions between any of the variables (all ps>.11).

4.2.2.2. Word N+1

Effects of Word N+1 Type: While skipping probability was unaffected by word n+1 type (both Fs<1), there was some indication that first fixations tended to fall further into the alternative high frequency words than the determiners, although this was an effect clearly not typical for all items (1.47

vs 1.29 character spaces: $F1(1,56)=8.45$; $p<.01$; $F2(1,40)=1.32$; $p=.26$). As can be seen from Table 4.2, individual fixation durations were, on average, marginally higher when word n+1 was an alternative high frequency word than when it was a determiner. This trend however was rather variable across subjects (first fixation duration: 246ms vs. 239ms: $F1(1,56)=3.13$; $p=.08$; $F2(1,40)=6.84$; $p<.05$; single fixation durations: 247ms vs 240ms: $F1(1,56)=2.27$; $p=.13$; $F2(1,40)=5.40$; $p<.05$ and last fixation duration: 242ms vs 237ms: $F1(1,56)=2.04$; $p=.15$; $F2(1,40)=3.72$; $p=.06$).

Table 4.2. *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for Word N+1.*

	Determiner				High Frequency Word			
	N+1 Identical		N+1 Invalid		N+1 Identical		N+1 Invalid	
	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid
First Fix	230	231	253	242	235	241	256	254
Last Fix	235	226	248	240	233	241	245	250
Single Fix	232	229	257	242	236	244	254	254
Gaze	116	131	151	141	121	146	146	141
Go-Past	126	144	173	173	129	159	166	162
Re-Reading	10	13	22	32	7	13	20	20
Skip Prob	51	45	42	44	51	42	47	47
Landing	1.10	1.28	1.33	1.47	1.40	1.36	1.47	1.60

There was a small increase in re-reading time for determiners compared to alternative high frequency words, however, this failed to achieve significance (19ms vs. 15ms; $F1(1,56)=1.97$; $p=.16$; $F2(1,40)=1.17$; $p=.29$). There was also no statistical evidence to suggest that word n+1 type influenced either of the remaining cumulative measures (gaze duration and go-past time: all $Fs<1$).

Effects of Word N+1 Preview: A standard word n+1 preview benefit was evident across all durational measures, with shorter inspection times when the preview was identical (first fixation duration: 234ms vs 251ms: $F1(1,56)=19.86$; $p<.001$; $F2(1,40)=13.14$; $p<.01$; last fixation duration: 234ms vs 246ms: $F1(1,56)=9.27$; $p<.01$; $F2(1,40)=5.29$; $p<.05$; and single fixation duration: 235ms vs 252ms : $F1(1,56)=17.39$; $p<.001$; $F2(1,40)=11.57$; $p<.01$; gaze duration: 129ms vs 145ms: $F1(1,56)=8.49$; $p<.01$;4$F2(1,40)=9.28$; $p<.01$; go-past time: 140ms vs 168ms: $F1(1,56)=16.63$; $p<.001$; $F2(1,40)=12.75$; $p<.01$; and first-pass re-reading time: 11ms vs 24ms: $F1(1,56)=12.86$; $p<.01$; $F2(1,40)=8.08$; $p<.01$). An invalid preview of word n+1 also encouraged first fixations to fall further into the word (1.47 vs 1.29 character spaces: $F1(1,56)=5.06$; $p<.05$; $F2(1,40)=3.74$; $p=.06$). Importantly, like many prior studies (e.g., Balota, Pollatsek & Rayner, 1985; Inhoff, 1989; Rayner, 1975; Starr & Inhoff, 2004; Rayner, 1975; White, Rayner & Liversedge, 2005), all durational measures show clear evidence for a word n+1 preview benefit.

Overall, skipping probability was reduced when word n+1 had received an invalid preview; although this was significant by items, this trend was clearly not typical for all subjects ($F1(1,56)=3.20$; $p=.08$; $F2(1,40)=4.40$; $p<.05$). There was, however, some indication that this might vary between the two word types with a trend towards an interaction ($F1(1,56)=3.00$; $p=.09$; $F2(1,40)=3.17$; $p=.08$). While there was a 5% reduced probability of skipping word n+1 when it had received an invalid preview in the determiner condition

(43% vs 48%: $F1(1,56)=4.04$; $p<.05$; $F2(1,40)=4.12$; $p<.05$), there was no difference in the case of the alternative high frequency words (46% vs 47%: both $Fs<1$). It appears that determiners are more likely to be skipped, but of course, only if they can be seen in the preview.

Effects of Word N+2 Preview: There was no evidence to suggest that word n+2 preview influenced first, last or single fixation durations on word n+1 (all $Fs<1$). The cumulative measures revealed a general pattern of increased inspection time when word n+2 had received an invalid preview, but effect sizes were small and unreliable (7ms in gaze duration: $F1(1,56)=1.30$; $p=.27$; $F2(1,40)=1.69$; $p=.20$; 12ms in go-past time: $F1(1,56)=2.19$; $p=.14$; $F2(1,40)=2.95$; $p=.09$; and 5ms in re-reading time: $F1(1,56)=1.62$; $p=.21$; $F2(1,40)=2.03$; $p=.16$).

There was some indication that an invalid preview of word n+2 encouraged first fixations to land further into word n+1, although this tendency clearly did not generalize across items (1.47 vs 1.29: $F1(1,56)=3.06$; $p=.08$; $F2(1,40)=1.01$; $p=.30$). There was, however, some reduction in the skipping of word n+1 when word n+2 had received an invalid preview ($F1(1,56)=3.20$; $p=.08$; $F2(1,40)=4.40$; $p<.05$), but as can be seen in Figure 4.4, there was also an interaction involving both word n+1 and word n+2 preview ($F1(1,56)=5.64$; $p<.05$; $F2(1,40)=5.18$; $p<.05$) with the skipping effect confined to cases in which word n+1 was parafoveally visible while fixating word n

($F1(1,56)=9.24$; $p<.01$; $F2(1,40)=8.61$; $p<.01$) and no effect when word n+1 had been invalid (Fs<1).

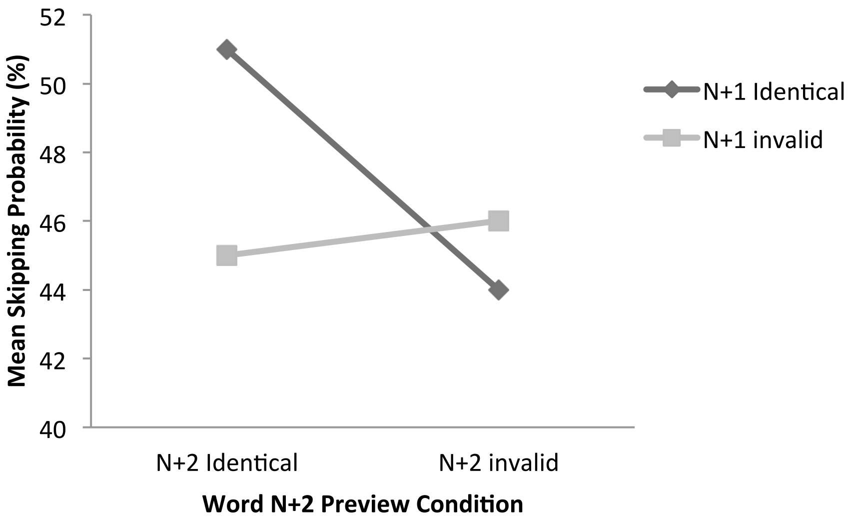


Figure 4.4. *Mean Skipping Probability (%) on Word N+1 as a Function of Word N+2 Preview for Each of the Two Word N+1 preview Conditions*

The direction of this effect lies in direct contrast to that shown on word n; leading to two potential interpretations. First, it may be that different responses to word n+2 illegality were engaged depending upon when that illegality was first detected. Specifically, something peculiar in a very remote location might attract attention (i.e., if detected on word n-1 an increased probability of skipping word n), whereas, if detected later, a more cautious reading strategy is adopted (i.e., if detected on word n an increased probability of fixating word n+1).

Alternatively, it seems more likely that this pattern simply reflects a trade-off. Since the preferred saccade travels, on average, 7-9 characters while reading English (McConkie, Kerr, Reddix & Zola, 1988), the condition associated with an increase in skipping probability on word n should also be associated with an increase in fixation probability of word n+1 – as seen here.

Since the cumulative measures in these analyses attribute a skipped word a duration of 0ms, it is perhaps not surprising that the corresponding interactions between the two preview manipulations were also present in gaze duration ($F1(1,56)=8.19$; $p<.01$; $F2(1,40)=6.66$; $p<.05$) and showed some trend in go-past ($F1(1,56)=3.51$; $p=.06$; $F2(1,40)=3.44$; $p=.07$); these interactions are displayed in Figures 4.5.

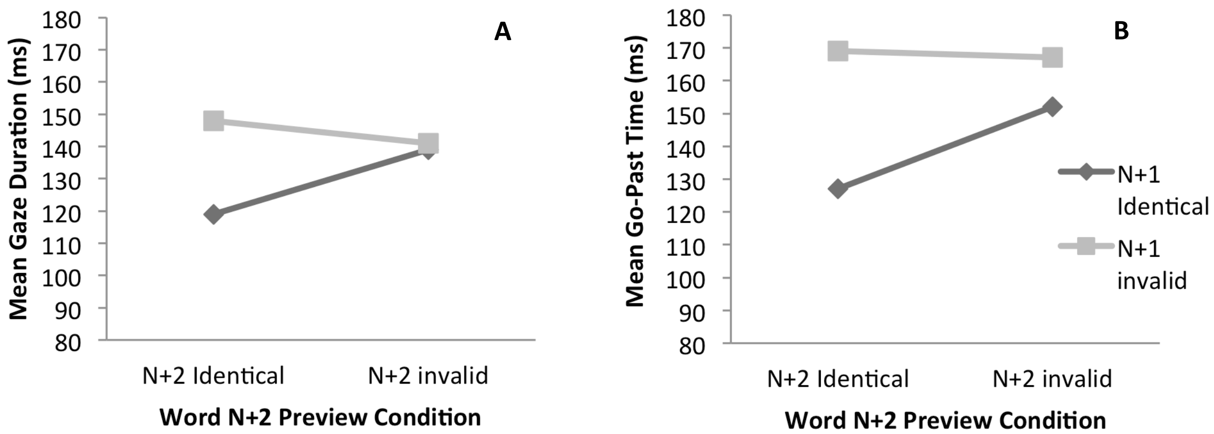


Figure 4.5. *Mean (A) Gaze Duration and (B) Go-Past Time (ms) on Word N+1 for Both Types of N+1 Preview as a Function of Word N+2 Preview.*

As the figures Illustrate, word n+2 preview did not influence gaze or go-past times on word n+1 when it had received an invalid preview (all $Fs<1$), however, when word n+1 had received an identical preview, an invalid preview of word n+2 significantly increased both gaze duration ($F1(1,56)=8.92$; $p=.01$; $F2(1,40)=7.04$; $p<.05$) and go-past time ($F1(1,56)=8.80$; $p<.01$; $F2(1,40)=7.03$; $p<.05$). While a similar pattern emerged in first fixation duration, the qualifying interaction was clearly not significant by-subjects ($F1(1,56)=1.74$; $p=.19$; $F1(1,40)=3.83$; $p=.05$). Since these interactions were confined to gaze duration and go past time, it seems that they were, at least primarily, the result of attributing a skipped word a value of 0ms. Consistent with this, the interaction disappeared for these measures when zeroes were excluded from the calculations (gaze duration: $F1<1.1$; $F2(1,40)=2.14$; $p=.14$ and go-past time: $Fs<1$).

No other interactions were significant on word n+1 (all $ps>.10$). It therefore appears clear that the effects of n+2 preview shown on word n+1 were driven primarily by differences in skipping probability. Since this family of effects is restricted to cases where word n+1 was valid, it is possible that their occurrence could be explained by the E-Z Reader model via a double attention shift. This possibility will be considered further in the Discussion.

4.2.2.3. Word N+2

Effects of Word N+1 Type: First fixations fell further to the right in word n+2 following a determiner rather than an alternative high frequency word (2.81 vs. 2.42 character spaces: F1(1,56)=23.69; p<.001; F2(1,40)=29.75; p<.001). There was also a 2% increase in the probability of skipping word n+2 if the preceding word was a determiner (9% vs. 7%: F1(1,56)=5.58; p<.05; F2(1,40)=5.55; p<.05). Since these two classes of words had been skipped with equal frequency (see above), these effects on word n+2 suggest an increase in word n+2 pre-processing when the preceding word had been a determiner than when it was an alternative high frequency word.

Consistent with this, inspection times were significantly shorter when the preceding word was a determiner (first fixation duration: 251ms vs. 262ms: F1(1,56)=16.91; p<.001; F2(1,40)=11.10; p<.01; last fixation duration: 242ms vs. 250ms: F1(1,56)=8.04; p<.01; F2(1,40)=3.82; p=.05; single fixation duration: 252ms vs. 267ms: F1(1,56)=15.25; p<.001; F2(1,40)=10.84; p<.01; gaze duration: 268ms vs. 299ms: F1(1,56)=53.31; p<.001; F2(1,40)=42.09; p<.001, and go-past time: 317ms vs. 359ms: F1(1,56)=35.22; p<.001; F2(1,40)=39.16; p<.001). The same trend was also observed in first-pass re-reading time (49ms vs. 60ms: F1(1,56)=2.97; p=.09; F2(1,40)=3.80; p=.06).

While the E-Z Reader model can simulate spillover effects, it is unclear whether the model is capable of simulating the results obtained in this study.

As will be recalled from Chapter 1, since L1 is a fixed proportion of L2 and saccadic programming takes a fixed time, 'spillover' effects reflect modulation in preview benefit dependent on the difficulty of the previously fixated word, with attention proceeding to the parafoveal word faster if the foveal word is 'easy' rather than 'difficult'. Thus differences in foveal word difficulty are a necessary prerequisite for spillover effects. There was, however, little reliable evidence to suggest that readers found the alternative high frequency words any more difficult to process than determiners in this case. The largest effect on word n+1 was a 7ms increase in single fixation duration while fixating word n+1 (reliable by items but not by subjects; see above) and no evidence at all in the cumulative measures. It is therefore difficult to see how this lack of difference on word n+1 translates into the rather sizable, highly significant, spillover effects observed on word n+2.

Of course, these effects could also reflect higher levels of integration difficulty with the alternative high frequency words than with the determiners. According to E-Z Reader 10, an integration difficulty could result in either longer durations on word n+2 or regressions out of word n+2 - both of which can be seen here. However, these effects are seen as a response to 'integration failure' and it is difficult to see how the integration of a short common adjective with the following noun would be likely to fail. Rather, it seems more likely that this reflects not a failure, but simply the complexity of higher level integration, with modified noun phrases more difficult to interpret than simple unmodified determiner-noun pairs.

Table 4.3. *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for Word N+2.*

	Determiner				**High Frequency Word**			
	N+1 Identical		**N+1 Invalid**		**N+1 Identical**		**N+1 Invalid**	
	N+2 Identical	**N+2 Invalid**	**N+2 Identical**	**N+2 Invalid**	**N+2 Identical**	**N+2 Invalid**	**N+2 Identical**	**N+2 Invalid**
First Fix	252	257	247	248	260	263	261	264
Last Fix	245	246	241	236	248	252	252	248
Single Fix	255	259	248	247	265	274	262	266
Gaze	262	263	275	272	297	292	305	301
Go-Past	293	294	330	348	348	335	373	379
Re-Reading	31	32	55	77	51	43	68	78
Skip Prob	9	10	9	7	6	9	6	7
Landing	3.05	3.12	2.58	2.50	2.56	2.48	2.23	2.43

Effects of Word N+1 Preview: While skipping rates were low and unaffected by word n+1 preview ($F1(1,56)=1.20$; $p=.28$; $F2(1,40)=1.38$; $p=.25$), earlier first landing positions were observed following a previously invalid preview of word n+1 (2.43 vs. 2.80 character spaces: ($F1(1,56)=26.14$; $p<.001$;

$F2(1,40)=51.51$; $p<.001$). An interaction between word n+1 type and word n+1 preview ($F1(1,56)=8.66$; $p<.01$; $F2(1,40)=9.25$; $p<.01$) suggests that this effect was strongest following a determiner (2.54 vs. 3.09 character spaces: $F1(1,56)=32.96$; $p<.001$; $F2(1,40)=45.99$; $p<.001$), although it was clearly still present following an alternative high frequency word (2.31 vs. 2.47 character spaces: $F1(1,56)=4.47$; $p<.05$; $F2(1,40)=4.42$; $p<.05$). It seems likely that this effect was driven by the adoption of a more cautious reading style following a previously invalid word n+1.

While word n+1 preview did not influence first, last or single fixation durations (both $Fs<1.1$; $F1(1,56)=1.62$; $p=.21$; $F2<1$; and $F1(1,56)=2.60$; $p=.11$; $F2(1,40)=1.43$; $p=.24$, respectively), there was a trend towards an invalid preview of word n+1 resulting in an increase in gaze duration on word n+2 (279ms vs. 288ms: $F1(1,56)=3.03$; $p=.08$; $F2(1,40)=3.32$; $p=.07$). This effect became significant in go-past time (318ms vs. 358ms: $F1(1,56)=30.47$; $p<.001$; $F2(1,40)=35.70$; $p<.001$), presumably because of less time spent re-reading when word n+1 had received a valid preview (39ms vs. 69ms: $F1(1,56)=33.93$; $p<.001$; $F2(1,40)=37.34$; $p<.001$).

A Spillover effect such as this is difficult to reconcile with the E-Z Reader model, since a saccade into word n+2 should not have been programmed until L1 processing of word n+1 was complete and the orthographic (and phonological) code had been extracted. Thus, while E-Z Reader can account for some spillover effects, these should not be orthographic in nature – as we see here. However, proponents of the serial perspective might argue that this effect could have resulted from a mislocated fixation with word n+1 accidently skipped – it was, after all, a short word – to which the immediate response might be mixed. First, a stay and process strategy could have contributed towards the trend in gaze, while a quick error correcting saccade back to word n+1 could account for the effects in first pass re-reading time and go-past time; this latter response is consistent with Nuthmann et al (2005; 2007) who implicate such a response in the IOVP. But while a mislocated fixation could conceivably give rise to such a pattern of effects, a trend towards inflated inspection times continued into the spillover region, where a mislocated fixation should be far less likely (see below).

Effects of Word N+2 Preview: Word n+2 preview did not influence inspection times on word n+2 when it was eventually fixated (first: $F1(1,56)=1.61$; $p=.21$; $F2<1$; last: $Fs<1$ and single fixation duration: $F1(1,56)=1.79$; $p=.18$; $F2<1$; gaze duration and go-past time: all $Fs<1$; first-pass re-reading: both $Fs<1.2$), nor did it influence first landing position or the probability of skipping (both $Fs<1$).

As can be seen from Table 4.3, if word n+1 had been a determiner, there was some indication that an invalid preview of word n+2 resulted in later first landing positions if word n+1 had received an identical preview, but earlier first landing positions if word n+1 had received an invalid preview. The opposite pattern occurred for cases in which word n+1 was an alternative high frequency word. The corresponding three-way interaction was not however significant ($F1(1,56)=2.53$; $p=.11$; $F2(1,40)=3.58$; $p=.06$). Indeed, there was no statistical evidence to suggest that word n+1 preview modulated the word n+2 preview effect when word n+1 was either a determiner (both $Fs<1.2$) or an alternative high frequency word ($F1(1,56)=2.22$; $p=.14$; $F2(1,40)=2.63$; $p=.11$).

No other interactions were significant (all $ps>.13$), and it is therefore clear that word n+2 preview did not influence any reading measure localized to word n+2.

4.2.2.4. Spillover Region

Effects of Word N+1 Type: Consistent with the increased probability of skipping word n+2 when it was preceded by a determiner compared to an alternative high frequency word, there was some indication that the first landing position in the spillover region was earlier for the former than the latter. This trend was, however, only marginally significant by-subjects (4.27 vs. 4.44 character spaces: $F1(1,56)=2.90$; $p=.09$; $F2(1,40)=1.98$; $p=.16$).

Table 4.4. *Fixation Time Measures (ms) and First Landing Positions (character spaces) for the Spillover Region.*

	Determiner				High Frequency Word			
	N+1 Identical		N+1 Invalid		N+1 Identical		N+1 Invalid	
	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid
First Fix	243	245	246	249	238	243	254	241
Gaze	495	499	483	493	496	502	502	481
Go-Past	560	550	557	555	565	580	576	561
Re-Reading	65	51	74	62	69	78	74	80
Landing	4.26	4.35	4.32	4.15	4.47	4.47	4.53	4.30

There was some evidence of a delayed 'cost' for an alternative high frequency word preceding the noun continuing into the spillover region. There was a trend in go-past time by-subjects (570ms vs 556ms: $F1(1,56)=3.45$; $p=.07$; $F2(1,40)=2.21$; $p=.14$), apparently driven by a numerical but non-significant trend in re-reading time (75ms vs. 63ms: $F1(1,56)=2.36$; $p=.13$; $F2(1,40)=1.81$; $p=.18$); with no evidence for a continuing effect in gaze duration (both $Fs<1$). Again, these results are consistent with the suggestion that there was more difficulty integrating an alternative high frequency word with the noun than integrating it with a determiner and that this difference either spilled over into the following region or made that noun phrase more difficult to integrate with the content that followed. These results suggest that inspection times can sometimes be as much influenced by integration factors as by lexical difficulty - something that is not currently well implemented in any model of eye movement control during reading.

Effects of Word N+1 Preview: The word n+1 preview spillover effect apparent on word n+2 continued into the first fixation duration of the spillover region, with longer durations following a previously invalid preview; but while this increase was significant by-subjects it was apparently not consistent over items (242ms vs. 247ms: $F1(1,56)=5.66$; $p<.05$; $F2(1,40)=2.51$; $p=.12$). This trend was restricted to first fixation duration only (gaze duration: $F1(1,56)=1.34$; $p=.25$; $F2(1,40)=1.71$; $p=.19$; go-past and re-reading time: all $Fs<1$). Again, it is worth noting that a spillover effect of word n+1 preview onto subsequent words to the right of word n+1 is extremely difficult to reconcile with the E-Z Reader model, especially here, where it spills over onto non-adjacent words making a mislocated fixation argument far less plausible.

Effects of Word N+2 Preview: Finally, there were no delayed main effects of word n+2 preview in the spillover region (all $Fs<1$). First fixation duration did, however, show some tendency towards a 3-way interaction between the two preview manipulations and word n+1 type ($F1(1,56)=3.03$; $p=.08$; $F2(1,40)=3.80$; $p=.06$) and a marginal 2-way interaction between the two preview manipulations ($F1(1,56)=2.51$; $p=.11$; $F2(1,40)=5.66$; $p<.05$). As can be seen in Table 4.4, while there was no apparent interaction in the effect of the two preview manipulations when word n+1 was a determiner (both

Fs<1), when it was an alternative high frequency word, a significant interaction between the two preview manipulations was apparent (F1(1,56)=5.77; p<.05; F1(1,40)=10.13; p<.01). Pairwise comparisons revealed that when word n+1 preview had been valid, an invalid preview of word n+2 increased first fixation duration by 5ms, although this was clearly nonsignificant by-subjects (F1(1,56)=1.29; p=.26; F2(1,40)=3.76; p=.06); however, when both words n+1 and n+2 had received invalid previews, first fixation duration was significantly shorter in the spillover region than when only word n+1 had received an invalid preview (F1(1,56)=4.64; p<.05; F2(1,40)=5.87; p<.05). It is acknowledged that these effects are somewhat tenuous, especially since they are only reflected in first fixation duration and there was no indication of a similar pattern on word n+2. They do, however, provide some evidence of a word n+2 preview benefit at a point after which word n+2 had been fixated.

There were no other interactions between any of the three variables within this region (all ps>.18).

4.2.3. General Discussion of Experiment 2

Evidence pertaining to word n+2 preview benefit has been mixed (e.g., Angele et al, 2008; Angele & Rayner, 2011; Kliegl et al, 2007; Radach et al, 2007; Radach et al, 2013; Rayner et al, 2007; Risse & Kliegl, 2011). The present study sought to investigate whether these differences may have been driven – at least in part – by variations in critical word lengths. Since evidence suggests that foveal word length modulates the expression of parafoveal-on-foveal effects (e.g., Hyönä & Bertram, 2004; Kennedy, Pynte & Ducrot, 2002), it is conceivable that variations in word n length might also be responsible for the inconsistency of results found in word n+2 preview experiments. It could be the case that the lengths of the critical words (n, n+1 and n+2) in some previous studies might have inhibited word n+2 pre-processing, potentially explaining these null results (e.g., Rayner et al, 2007; Angele et al, 2008; Angele & Rayner, 2011).

Despite employing items here that should, according to Radach et al (2007), optimally allow for word n+2 preview, only a limited set of word n+2 preview effects materialized. Importantly, there was no evidence of a word n+2 preview effect on the inspection of word n+2, and while there was a numerical trend towards increased inspection times on word n+1 when word n+2 had received an invalid preview, these so-called delayed parafoveal-on-foveal effects (Kliegl et al, 2007) were unreliable. In fact, the only reliable evidence for a word n+2 preview effect after the boundary had been crossed was confined to first fixation duration in the spillover region, with an invalid preview of word n+2 resulting in increased inspection times, but only when word n+1 was an alternative high frequency word that had received an invalid preview while fixating word n. Given the absence of word n+2 preview benefit on word n+2, this effect is difficult to interpret and should be treated with caution. It seems doubtful therefore, that variability in critical word length alone can account for the history of discrepant results in word n+2 preview experiments.

Despite the lack of post-boundary effects, there was evidence that word n+2 preview modulated fixation durations prior to crossing the invisible boundary. The present n+2 parafoveal-on-foveal effect mirrors that reported by Kliegl et al (2007), in that it was only present when word n+1 was an alternative high frequency word, rather than a determiner. It was, however, expressed in a later measure than found by Kliegl et al (go-past rather than gaze) and fell in the opposite direction, with shorter durations when word n+2 received a nonword preview. These inconsistencies will be returned to below, but first, let us consider why the effects might only be present when word n+1 was an alternative high frequency word. As Kliegl et al acknowledge, this pattern is surprising since it is contrary to the prediction of the word-grouping hypothesis (Radach, 1998). According to this theory, there is a preference to process determiner-noun pairs as one perceptual unit; with this preference limited to these word pairs and not apparent for non-determiner-noun pairs (c.f., Drieghe et al, 2008). Consequently, this would predict enhanced word n+2 pre-processing when word n+1 was a determiner, rather than an alternative high frequency word.

This effect seems less counterintuitive, however, if one considers that the alternative high frequency words in the present study carry meaning (e.g., “one”, “two”, “old”, “new”, “his”, “her”), but the interpretation of this meaning very much depends upon the nature of the subsequent noun. This interdependence between the two words is not present for determiner-noun pairings, since an article carries syntactic, rather than semantic information. Consequently, while determiners might be ‘easier’ to process, their interpretation does not depend upon the upcoming word, and so attention might not be encouraged to stretch to word n+2 in the same way as when word n+1 acts as a modifier. This explanation might initially appear to be inconsistent with the finding that first landing positions within word n+2 were shifted to the right and word n+2 was more likely to be skipped following a determiner compared with an alternative high frequency word. However, if during fixation of word n+1, both word n+1 and n+2 were being processed in parallel, and since determiners are ‘easier’ to process, then more word n+2 pre-processing should have been possible. Attention may be distributed differently across word units and/or phrases depending on where the reader is fixating.

As will be recalled from Chapter 2, it is still unclear why some parafoveal-on-foveal effects appear to be expressed in an orthodox direction (e.g., Kliegl et al, 2007) while others are expressed in a reversed direction (as in the present study). One potential explanation for this difference in response is that word n length varied considerably in Kliegl et al (4-13 letters), while it was held constant in the present study (all items except 1 had 6-letters). Previous studies have indeed indicated that differences in foveal word length may contribute towards such inconsistencies (Hyönä & Bertram, 2004; Kennedy & Pynte, 2005; Kennedy, Pynte & Ducrot, 2002). Also, some authors have suggested the possibility that the inconsistences might have been due to different writing systems (Kennedy & Pynte, 2005; Risse & Kliegl, 2011); a dimension on which the present study and Kliegl et al also differed. Finally, this variation in response might be related to the frequencies of the non-function words employed in each of these studies. While this word was relatively low frequency in Kliegl et al's study (mean 29/million), it was always a very high frequency word in the present study (mean 3040/million). As lower frequency words will constrain attention more than high frequency words, differences in word n+1 frequency might affect the reader's confidence in the perception of a parafoveal nonword. For example, high confidence in the presence of a high frequency intervening word might result in the reader deciding to press ahead and inhibit a regression, while low confidence, associated with a lower frequency intervening word, might encourage a more cautious reading strategy.

Word n+2 preview also influenced targeting decisions on both word n and n+1. Again, this result is not without precedence. Some studies have reported an increased probability of fixating word n+1 if word n+2 received an invalid preview (e.g., Pynte, Kennedy & Ducrot, 2004; Radach, 2007), while others report that invalid previews of word n+2 encouraged word n+1 skipping

(Angele et al, Exp 1, 2011; Radach et al, 2013). The present study obtained both patterns: an increased probability of skipping word n, but a reduced probability of skipping word n+1 when word n+2 had received an invalid preview. As discussed above, the skipping effect on word n+1 likely has its origins in the skipping effect on word n. This is the first time such an effect has been found as far 'upstream' as word n, and it is significant that a decision to skip word n must have originated from word n-1, potentially reflecting a word n+3 effect.

Proponents of the E-Z Reader model might argue that since the effects were confined to cases where word n+1 was available, they could stem from a triple attention shift. But while theoretically possible, such an explanation would push the model beyond its current capabilities. For example, Schotter et al (2014) recently demonstrated that while E-Z Reader can simulate word n+2 preview benefits - via a double-attention shift – the tight time constraints for these attention shifts prevent word n+2 from ever entering the L2 stage of lexical processing. It is unlikely therefore that the current version of the model would be able to accommodate the present set of results. Such a proposal also appears to undermine the assumption that there is a tight coupling between fixation location and the locus of attention during reading.

Alternatively, proponents might suggest that these effects were driven by the low level attentional scan detecting upcoming irregularity (e.g., Angele and Rayner, 2011; Reichle et al, 2003). However, the viability of such a mechanism, and whether it could account for the range of targeting based effects seen in this experiment and others, is difficult to assess without actual implementation in the model. Such a mechanism would, however, require a rather wide effective range in order for word n+3 irregularities to be detected.

This study provided further evidence of orthographic parafoveal-on-foveal effects on n (e.g., Angele et al; 2011; Rayner, 1975; Starr and Inhoff, 2004). As the trends and effects were apparent in gaze duration and go-past time, rather than individual fixation durations, this suggests they were not driven by mislocated fixations aimed at word n+1 as Drieghe et al (2008) might suggest. An undershoot of word n+1 might arise for two reasons: an accidental refixation of word n or a failed skip. Any refixations of word n should, however, typically overshoot not undershoot word n given that the preferred saccadic length (7-letters; McConkie et al, 1988) is longer than word n in the present experiment (6-letters). The experimental materials of this study therefore make it unlikely that a mislocated refixation would arise on word n often enough to be driving this effect. The fact that word n received, on average, just 1.17 fixations appear to corroborate this reasoning. If it is argued that the effect might be driven by a failed skip of word n – resulting in an erroneous fixation of word n followed by a stay and process response – it is difficult to see why there should only be evidence for the effect in the cumulative measures and not in single or first fixation durations.

It is, however, easier to envisage how a low level attentional scan might produce orthodox parafoveal-on-foveal effects of word n+1 preview on word n – as reported here – than to understand why and how it might produce differential targeting strategies based on word n+2 preview. Therefore, it could be argued that this might be driving the effect. Again, simulations seem essential in order to understand whether the E-Z Reader model would actually be capable of simulating such effects without jeopardizing its ability to account for the benchmark findings which it currently does a good job of explaining.

However, neither the low level attentional scan nor the mislocated fixations account of parafoveal-on-foveal effects provides a plausible explanation for why a determiner to the right of fixation should result in increased re-reading time than when it was an alternative high frequency word. As previously discussed, this effect could be accommodated within the E-Z Reader framework if one assumes that word n+1 was fully processed while fixating word n – via an attention shift – which resulted in increased integration failures for determiners compared to the alternative high frequency words, reflected in increased regressions. This argument only works, however, if the high level processing module is allowed to target a regression to a point earlier than where the processing difficulty was first detected. Additionally, as mentioned previously, it seems incongruent to assume that a determiner should be associated with more processing difficulty than an alternative high frequency word.

On word n+1, there was some evidence for a pattern of increased inspection time when it was an alternative high frequency word rather than a determiner, although not with statistical reliability. But while such word type effects often arise on the word itself (e.g., Angele & Rayner, 2011), this is not always the case (e.g., Risse & Kliegl, 2011). There was evidence here that an identical preview of a determiner was skipped more often than either a nonword or an alternative high frequency word. This aligns with previous research in demonstrating that determiners are skipped more frequently than other 3-letter words (e.g., Angele & Rayner, 2011; 2013; Dreighe et al, 2008; Just & Carpenter, 1983; O'Regan, 1979).

Word type effects on word n+1 are therefore not out of the ordinary. However, the rather sizable spillover effects of word n+1 type seen on words n+2 and in the spillover region do speak to the extent to which higher-level linguistic processing impacts the eye movement record. To some extent this is reflected in the addition of a higher-level processing module to E-Z Reader, although at present this is limited to dealing with integration failures rather than integration difficulty *per se*. The present set of results clearly highlight the importance of incorporating these and other aspects of high-level processing difficulty within models of eye movement control during reading.

It is not clear, however, how a delayed word n+1 preview spillover effect can be compatible with the model. A decision to plan a saccade out of word n+1 should coincide with the completion of L1 on word n+1 – that is, the extraction of orthographic and phonological information from that word – so effects of this nature should never spillover onto subsequent words. And while it could be argued that they arise following an accidental skip of word n+1, followed by a stay and process response, this seems unlikely to explain the trend still present in first fixation duration in the spillover region, since a mislocated fixation of this magnitude is highly improbable. It also appears unlikely that this effect can be accounted for by a spillover cost of skipping word n+1, since this should have resulted in increased durations following valid previews of a determiner (associated with the most skipping), not shorter durations – as seen here. Nor is this the first time that spillover effects of this nature have been observed. For example, Radach et al (2013) recently reported increased inspection times on word n+3 if word n+2 had received an invalid preview.

As discussed, many of the effects obtained in this study seem to be incompatible with the current version of the most prominent model of the serial variety: the E-Z Reader model 10. Little, however, has been said regarding whether SWIFT would be capable of accommodating the observed pattern of results. This model has previously demonstrated its ability to simulate word n+2 preview benefit on words n+1 and n+2 (Risse & Kliegl, 2014), but of course, effects of this nature were not found in the present study. Indeed, it could be argued that the lack of such effects - under what should have been optimal conditions - is detrimental to the parallel perspective in general and SWIFT specifically.

The SWIFT model should, however, cope relatively well with the finding that word n+1 orthographic parafoveal-on-foveal effects were only observed in cumulative measures. It will be recalled from Chapter 1 that the foveal inhibition mechanism within SWIFT can only inhibit saccades from the foveal, not parafoveal words. However, target selection effects could re-produce these orthographic parafoveal-on-foveal effects. Specifically, the probability that word n will be refixated depends upon whether it has the highest level of activation at the point at which a saccade is committed to action. Therefore, differences in parafoveal word n+1 difficulty influence word n refixation probability, allowing word n+1 parafoveal-on-foveal effects to be expressed in cumulative measures on word n. The same process could potentially produce (weaker) word n+2 parafoveal-on-foveal effects (Risse, Engbert & Kliegl, 2008), although it seems likely that the current version of SWIFT would fail to capture the dynamics of the present word n+2 preview effects and their interactions with word n+1 type and preview.

It is also unlikely that SWIFT would be capable of reproducing the differences in targeting strategies seen on words n and n+1, apparently driven by word n+2 preview. Since the rate of lexical processing in the model is dependent primarily upon the spatial location of the word’s individual letters within the perceptual span[24], it is difficult to see how a word 3-words upstream could receive enough pre-processing for a difference in lexical activation to trigger differences in targeting strategy.

It appears therefore, that these results present a challenge to current versions of both parallel and serial models. However, it should be emphasized that while both model types would fail to capture the complexities of the present dataset, the results, overall, do seem to fit more parsimoniously within a class of models that permits the lexical processing of multiple words in an overlapping fashion.

4.3. EXPERIMENT 3

Experiment 2 tested the range over which word n+2 preview benefit might occur. The majority of recent research into word n+2 preview has focused on keeping word n+1 short, typically just three characters (e.g., Kliegl et al, 2007; Radach et al, 2007; Risse et al, 2011; Angele & Rayner, 2011, Radach et al, 2013). From a parallel perspective, this makes sense, since it will increase the probability that word n+2 will fall within the effective span of apprehension while fixating word n. However, parallel models such as SWIFT should also be capable of reflecting word n+2 preview benefit over a longer range, albeit to a lesser extent as a consequence of acuity constraints. In contrast to parallel models, the mechanisms within serial models, such as E-Z Reader, should

[24] This can, in SWIFT-3, be modulated by foveal load (Schad & Engbert, 2012).

typically prevent word n+2 preview benefit from accruing over longer eccentricities. Proponents of the serial perspective maintain that E-Z Reader is capable of accounting for word n+2 preview benefit via double attention shifts – a process with an increased probability of occurrence with a shorter word n+1. Finding word n+2 preview benefit over a longer range would therefore provide a greater challenge for the E-Z Reader model and consequently provide a platform on which the two models could more easily be differentiated.

Testing for the presence of word n+2 preview benefit over a longer range should also speak to the plausibility of the mislocated fixation account of so-called ‘delayed parafoveal-on-foveal effects’ (Kliegl et al, 2007) in which inspection times on word n+1 increase as a consequence of word n+2 being presented as an invalid nonword preview while fixating word n. This account of delayed parafoveal-on-foveal effects was proposed by proponents of the serial perspective (e.g., Schotter, Angele & Rayner, 2012) to explain such effects within the serial framework of the E-Z Reader model. Specifically, it is argued that occasionally a double attention shift occurs, which triggers a new saccadic program to be initiated to take the eye to word n+2. However, due to oculomotor error, the saccade falls short of word n+2 and instead fixates word n+1, from where word n+2 is then processed. Because word n+2 had received some preview benefit via the double attention shift while fixating word n, this results in the word n+2 preview benefit being expressed on word n+1 instead of on word n+2.

Since fixation probability increases with word length (Rayner & McConkie, 1976), employing longer words in the position of word n+1 should reduce skipping probability and in so doing, produce fewer occasions in which a failed skip could routinely result in word n+2 effects being found on word n+1. Obtaining delayed parafoveal-on-foveal effects over longer ranges would therefore call into question the plausibility of the argument that these effects can be the product of a failed skip coupled with a stay and process response. Given that such an explanation is critical for the E-Z Reader model, employing longer words in the position of n+1 will again provide a platform on which the serial and parallel models can more easily be differentiated.

Finally, increasing the length of word n+1 will help determine whether word n+2 preview effects are restricted to cases where word n+1 contains 3-letters. This is important since there is some evidence that very short words, such as 3-letter words, are treated differently from words of a longer length. Angele and Rayner (2013) specifically argued this in the case of determiner previews, which were skipped more frequently than either nonword previews or identical verb previews, even though the determiner violated the syntactic constraints of the sentence. It appears that different strategies and/or attentional mechanisms might be engaged with determiners (and potentially also other very short words), making them perhaps not the ideal intervening word for investigating the typicality of parallel processing. If parallel processing is a usual process, then it should be generalisable to a variety of types of word and not simply restricted to determiners or other very high frequency 3-letter words. This is especially important since some of the research showing word n+2 preview effects has exclusively employed determiners in the position of word n+1 (e.g., Radach et al, 2013). More generally, it is possible that since 3 letter words are skipped so frequently, different attentional mechanism might be engaged with them, so again, it is important to test whether these effects occur over a longer range where word n+1 skipping should not be such a factor.

Experiment 3 therefore tested word n+2 preview benefit with an intervening word n+1 which was either 4- or 6-letters in length. Longer words have been tested in the position of word n+1 in the past, but there were factors in these experiments that might have contributed to their null results. For example, as will be recalled from Chapter 2, Rayner et al (2007) used 4 letter words in the position of word n+1, but in that experiment word n+1 frequency tended to be low. And while Angele et al (2008) also tested longer word n+1s, their length was quite variable, ranging from 4- to 10-letters, with results from the longer words perhaps masking any effects present with the shorter words. This study therefore investigated differential effects of 4- and 6-letter words separately, and to prevent word n+1 from absorbing a high level of attentional resources, all critical words – including word n+1 – were of high frequency.

To summarise: if attention is allocated in a strictly serial sequential fashion during reading, then longer words in the position of word n+1 should prevent routine double attention shifts from occurring and therefore inhibit the expression of word n+2 preview effects. Given the reduced probably of skipping longer intervening words, there should also be no evidence of delayed parafoveal-on-foveal effects, especially when the intervening word contains 6-letters.

While the longer word n+1s should reduce the probability of a double attention shift, a second control was put in place to prevent this from occurring: as in Experiment 2, word n+1 also received either an identical or an invalid preview. An invalid preview of word n+1 should remove the conditions in which a double attention shift can, according to the E-Z Reader model, legitimately take place.

If, however, attention is distributed in a parallel fashion during reading, the present set of conditions should not prevent word n+2 preview effects occurring on words n, n+1 or n+2. They might, however, be somewhat attenuated with the inclusion of a longer word n+1 since this will reduce the processing efficacy of words n+1 and n+2 as a consequence of acuity constraints.

4.3.1. Method

4.3.1.1. Participants

Sixty-four native English speakers with normal or corrected-to-normal vision and with no known reading difficulties took part in the experiment. Each received course credit for their participation.

4.3.1.2. Materials and Design

Forty-eight experimental items were constructed. Each sentence frame contained an initial noun phrase followed by a 6-letter verb – designated as word n. This was followed by either a four- or six-letter word (n+1) and then by a 6-letter noun (n+2). To encourage pre-processing of word n+2 while fixating word n, words n, n+1 and n+2 typically formed verb-adjective-noun triplets. For the same reason, these words were also high frequency (word n: M=87, SD=28; word n+1: 4-letters: M=436, SD=542; word n+1: 6-letters, M=152, SD=175; word n+2: M=147, SD=241; all values per million; estimated using the Kuçera and Francis (1967) norms). Although it was not possible to construct items with no overall frequency difference between the 4 and 6 letter adjectives, average word frequencies were kept very high – over 150/million – and as Murray and Forster (2004) discuss – there is very little modulation of lexical access time by word frequency above this point. Sentences ranged from 60 to 85 characters in length.

Similarly to Experiment 2 words n+1 and n+2 were initially presented in parafoveal vision in either their correct form or as a nonword of equivalent length – an 'invalid' preview. Nonwords typically formed orthographically illegal letter strings that matched their respective target word's word envelope. Thus there were 4 experimental preview conditions in total, with one, the other, neither or both of word n+1 and n+2 initially receiving an invalid preview.

The boundary used to trigger the contingent change always occurred immediately following the verb (word n). Crossing this boundary triggered a display change to occur, presenting words n+1 and n+2 in their 'correct' form. As can be seen in Figure 4.6 below, this resulted in a 2 (word n+1 length) x2 (word n+1 preview) x2 (word n+2 preview) design with a total of 8 conditions. The full set of experimental items can be seen in Appendix H.

```
Four-Letter Word      n                   n+1   n+2

2a) The young child caught| (many leaves) many leaves as they fell
       from the old oak tree

2b) The young child caught| (many fonizc) many leaves as they fell
       from the old oak tree

2c) The young child caught| (esrg leaves) many leaves as they fell
       from the old oak tree

2d) The young child caught| (esrg fonizc) many leaves as they fell
       from the old oak tree

Six-Letter Word       n                     n+1   n+2

2e) The young child caught| (little leaves) little leaves as they fell
       from the old oak tree

2f) The young child caught| (little fonizc) little leaves as they fell
       from the old oak tree

2g) The young child caught| (huflka leaves) little leaves as they fell
       from the old oak tree

2h) The young child caught| (huflka fonizc) little leaves as they fell
       from the old oak tree
```

Figure 4.6. *Example item in each of the 4 parafoveal preview conditions for each of the two word length conditions (2a-2d: 4-letter words; 2e-2g: 6-letter words). Parafoveal previews are presented in parentheses, while target words n+1 and n+2 are underlined. The boundary location is denoted by the symbol: "|".*

Eight counterbalanced item files were constructed. Each participant experienced all preview conditions across an equal number of items, but saw only one version of each item. The particular allocations of items to files and subjects to files were treated as between-groups dummy variables in the following analyses.

To ensure normal reading for comprehension, 10% of the experimental items were followed by a comprehension question. In addition, a further 56 similar items with 12 comprehension questions were constructed as filler items[25]. Eight separate practice items preceded the experimental items; half of these were accompanied by a comprehension question.

4.3.1.3. Apparatus and Procedure

This was identical with Experiment 2.

4.3.2. Results and Discussion

For purposes of analysis, four zones were defined in each of the experimental items: one corresponding to each of the words n, n+1 and n+2 (zones 1, 2 and 3, respectively), and a 3-word 'spillover' region (zone 4). Fixations falling on the space preceding each of these were also considered to have fallen into the relevant region.

```
Zones:                 1      2      3         4
The young child| caught| many| leaves| as they fell| from the old oak
tree.
```

The same fixation time and saccadic measures employed in Experiment 2 were also analysed in this study.

[25] Forty-eight of these formed the experimental items for the first experiment reported in this chapter.

A repeated measures 2 (4- vs. 6-letter word) x 2 (word n+1 identical vs. invalid) x 2 (word n+2 identical vs. invalid) analysis of variance (ANOVA) was conducted for each of the measures for zones 1 to 4. Participants (F1) and Items (F2) were treated as random variables and file was treated as a between-groups dummy factor in both analyses.

Participants achieved an overall accuracy rate of 80%, suggesting they had read the sentences carefully.

4.3.2.1. Word N

Effects of Word N+1 Length As is apparent from Table 4.5, first landing position and skipping probability were both unaffected by word n+1 length (all $Fs<1$). Word n+1 length also did not modulate first fixation duration ($F1(1,56)=1.36$; $p=.25$; $F2(1,40)=2.75$; $p=.19$). All remaining durational measures showed a general pattern of increased inspection time when the word to the right was a 6- compared to a 4-letter word but these differences were small and clearly unreliable (single fixation duration: 255ms vs. 249ms: $F1(1,56)=2.89$; $p=.09$; $F2(1,40)=2.47$; $p=.12$; last fixation duration: 248ms vs. 243ms $F1(1,56)=1.78$; $p=.18$; $F2(1,40)=2.74$; $p=.10$; gaze duration: 262ms vs. 255ms: $F1(1,56)=2.08$; $p=.15$; $F2(1,40)=2.48$; $p=.12$; go-past time: 296ms vs. 286ms: $F1(1,56)=2.28$; $p=.13$; $F2(1,40)=3.05$; $p=.08$ and first-pass re-reading: 34ms vs. 31ms: both $Fs<1$).

While numerically small and unreliable, this pattern is nonetheless interesting since the durations fall in the opposite direction to the predictions of the E-Z Reader model. Specifically, there was a higher probability of skipping the parafoveal 4-letter compared to the parafoveal 6-letter words (see below) and since the mechanism responsible for word skipping in the model requires saccadic re-programming from word n+1 to word n+2 – a time-costly process – fixation durations preceding 4-letter words should have been higher, not shorter than for 6-letter words. While the differences were small and non-significant, the numerical trends go against the mechanism in E-Z Reader involved in word skipping. In contrast, these results do appear congruent with the findings of Engbert & Kliegl (2005) who report, for skipped parafoveal words, a positive correlation between parafoveal word length and pre-skip fixation durations.

Table 4.5 *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for Word N.*

	Four Letter Word				Six Letter Word			
	N+1 Identical		N+1 Invalid		N+1 Identical		N+1 Invalid	
	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid
First Fix	243	248	249	240	253	246	249	246
Last Fix	239	244	252	238	251	248	249	243
Single Fix	243	253	255	243	260	252	255	252
Gaze	251	257	262	252	261	263	256	267
Go-Past	301	285	283	274	295	288	292	309
Re-Reading	50	28	22	22	33	24	36	42
Skip Prob	12	12	12	12	12	10	13	10
Land Pos	2.98	2.86	2.81	2.80	2.87	2.80	2.89	2.88

Effects of Word N+1 Preview While there were no main effects of word n+1 preview on word n inspection time, first landing position or skipping rate (all Fs<1), there was a trend towards an interaction between word n+1 preview and word n+1 length in go-past time ($F1(1,56)=4.91$; $p<.05$; $F2(1,40)=3.18$; $p=.08$). It can be seen from Figure 4.7.A, that the length of the parafoveal nonword to the right of fixation influenced go-past time, with longer durations for longer parafoveal nonwords ($F1(1,56)=6.23$; $p<.05$; $F2(1,40)=5.89$; $p<.05$). This effects was clearly restricted to the nonword previews, with no significant length effects when the item in parafoveal vision was a word (Fs<1).

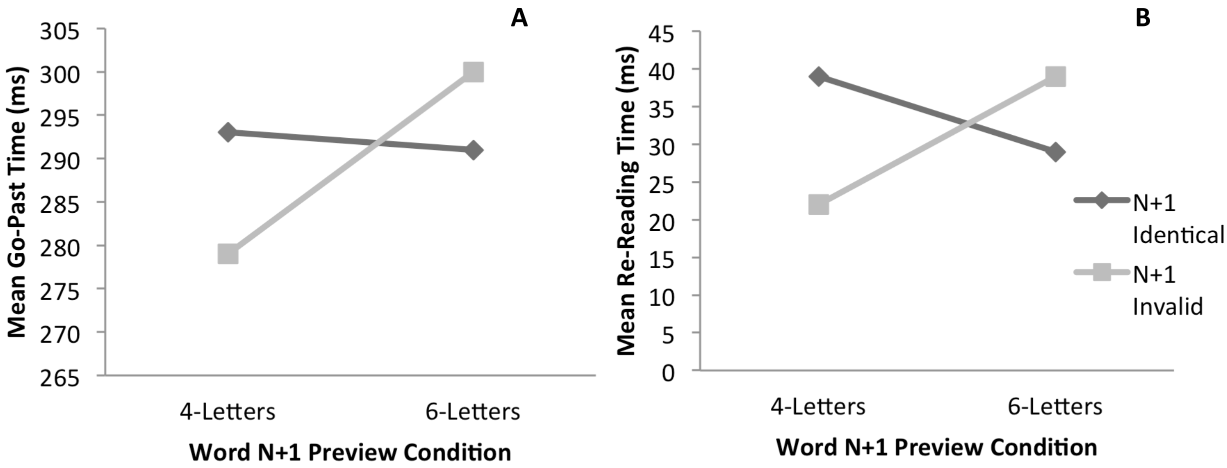


Figure 4.7. *Mean (A) Go-Past Time and (B) First-Pass Re-Reading Time (ms) on Word N for the Two Word N+1 Preview Conditions as a Function of Word N+1 Length*

Since go-past time includes time spent re-reading, it is unsurprising that the corresponding interaction was also apparent in first pass re-reading time ($F1(1,56)=7.70$; $p<.01$; $F2(1,40)=8.63$; $p<.01$). As shown in Figure 4.7.B, more time was spent re-reading when the parafoveal nonword was 6- rather than 4-letters in length ($F1(1,56)=5.24$; $p<.05$; $F2(1,40)=6.40$; $p<.05$). Again, no significant difference was observed when it was a word in parafoveal vision ($F1(1,56)=2.90$; $p=.09$; $F2(1,40)=1.71$; $p=.20$). These results highlight the importance of considering parafoveal word length when investigating parafoveal-on-foveal effects. It is clear that the length of the nonword to the right of fixation affected the expression of the parafoveal-on-foveal effect; therefore, in order to avoid the possibility that such differences might cancel one another out, future experiments should take care to ensure parafoveal word length remains constant, when parafoveal-on-foveal effects are investigated.

Effects of Word N+2 Preview There were no main effects of word n+2 preview on word n landing position (both $Fs<1$), skipping rates ($F1(1,56)=1.28$; $p=.26$; $F2(1,40)=1.99$; $p=.16$) or inspection times (first fixation duration: $F1(1,56)=1.43$; $p=.23$; $F2<1$; single fixation duration $Fs<1$; last fixation duration: $F1(1,56)=1.78$; $p=.18$; $F2(1,40)=1.12$; $p=.30$; gaze duration and go-past time $Fs<1$; first-pass re-reading time: $F1(1,56)=1.33$; $p=.25$; $F2(1,40)=1.33$; $p=.25$).

There was evidence, however, that first pass re-reading time was influenced by word n+2 preview, but that this was modulated by word n+1 availability ($F1(1,56)=3.84$; $p=.05$; $F2(1,40)=4.04$; $p<.05$). As can be seen in Figure 4.8, readers spent significantly more time re-reading when word n+2 was available, but only when word n+1 was also parafoveally available ($F1(1,56)=4.44$; $p<.05$; $F2(1,40)=5.85$; $p<.05$). The same pattern was not observed when word n+1 received an invalid preview (both $Fs<1$).

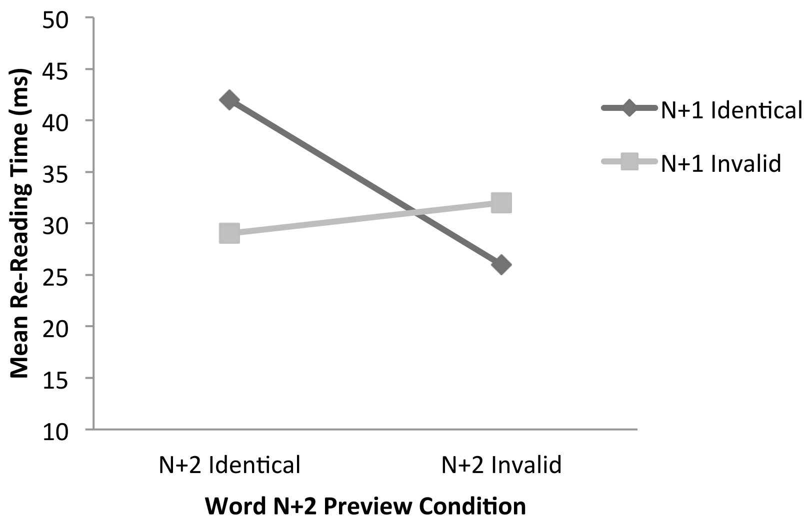


Figure 4.8. *Mean First-Pass Re-Reading Time (ms) on Word N for the Two Word N+1 Preview Conditions as a Function of Word N+2 Preview.*

This pattern appears to suggest that readers spent more time re-reading prior to moving forward if there was nothing peculiar in the parafovea. Proponents of serial models might suggest that this can be accounted for by a double attention shift, with word n+1 being identified while fixating word n, allowing attention to proceed to word n+2, thereby allowing it to influence reading strategy adopted while fixating word n. It should be borne in mind, however, that word n+1 in this experiment was not a 3-letter word that is routinely skipped – it had an average of 5-letters and was skipped just 20% of the time. How well the E-Z Reader model could account for a word n+2 parafoveal-on-foveal effect over this range will be returned to in the Discussion.

Another interaction between the two preview manipulations in last fixation duration does not, however, lend itself so naturally to an explanation in terms of a double attention shift. This interaction was significant by-subjects but only marginally significant by items (F1(1,56)=4.10; p<.05; F2(1,40)=3.22; p=.08); its form is shown in Figure 4.9.

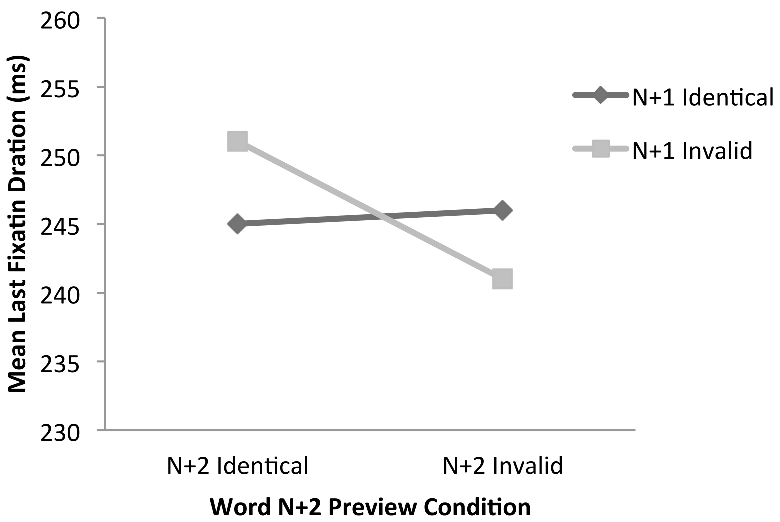


Figure 4.9. *Mean Last Fixation Duration (ms) on Word N for the Two Word N+1 Preview Conditions as a Function of Word N+2 Preview.*

As Figure 4.9 suggests, an invalid preview of word n+2 tended to reduce last fixation duration, but only when word n+1 was also parafoveally invalid (F1(1,56)=7.26; p<.01; F2(1,40)=3.49; p=.07; n+1 available: both Fs<1). Unlike the preceding interaction in first-pass re-reading time, this interaction cannot be explained within the E-Z Reader model by means of a double attention shift, since the invalid preview of word n+1 should have stalled the progression of attention to word n+2. Such effects appear to implicate a parallel extraction of information across several words, the range of which appears to stretch beyond that previously reported (e.g., Angele et al, 2008; Rayner et al, 2007). The plausibility of this effect being driven by a low level attentional scan, as suggested by proponents of the E-Z Reader model, will be returned to in the General Discussion.

While there was a trend towards a three-way interaction between the two preview manipulations and word n+1 length in single and first fixation duration, these interactions were clearly non-significant by-subjects and only marginal by-items (Single fixation duration: $F1(1,56)=2.18; p=.14$; $F2(1,40)=3.44; p=.07$ first fixation duration: $F1(1,56)=1.53$; $p=.22$; $F2(1,40)=2.89$; $p=.09$). No other interactions involving any combination of the three factors achieved significance either by-subjects or by-items (all $ps>.10$).

4.3.2.2. Word N+1

Effects of Word N+1 Length As is apparent from Table 4.6, inspection times tended to be longer on word n+1 when it was a 6- rather than a 4-letter word. This was not-significant in first fixation duration (262ms vs. 257ms: $F1(1,56)1.93$; $p=.17$; $F2(1,40)=3.48$; $p=.07$) approached significance in last fixation duration (261ms vs. 255ms: $F1(1,56)=3.33$; $p=.07$; $F21,40)=5.14$; $p<.01$) and was significant in all other durational measures (single fixation duration:

269ms vs. 259ms: $F1(1,56)=8.72$; $p<.01$; $F1(1,40)=9.90$; $p<.01$; gaze duration: 284ms vs. 192ms: $F1(1,56)=158.48$; $p<.001$; $F2(1,40)=275.44$; $p<.001$; go-past time: 325ms vs. 218ms: $F1(1,56)=131.54$; $p<.001$; $F2(1,40)=158.16$; $p<.001$; and first-pass re-reading time: 42ms vs. 27ms: $F1(1,56)=10.52$; $p<.01$; $F2(1,40)=6.81$; $p<.05$). Four-letter words were also skipped more frequency (30% vs. 9%: $F1(1,56)=100.67$; $p<.001$; $F2(1,40)=180$; $p<.001$) and, not surprisingly, were associated with earlier first landing positions than 6-letter words (1.99 vs. 2.88 character spaces: $F1(1,56)=152.54$; $p<.001$; $F2(1,40)=184.17$; $p<.001$).

Table 4.6. *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for Word N+1.*

	Four Letter Word				Six Letter Word			
	N+1 Identical		N+1 Invalid		N+1 Identical		N+1 Invalid	
	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid
First Fix	244	260	265	259	260	257	263	268
Last Fix	240	254	264	261	257	260	263	266
Single Fix	244	258	270	263	265	264	271	276
Gaze	169	184	210	205	276	263	303	292
Go-Past	198	207	237	230	317	293	352	339
Re-Reading	30	24	28	25	40	30	50	46
Skip Prob	34	33	25	28	9	12	5	11
Landing	1.83	2.07	2.06	2.00	2.98	2.81	2.85	2.85

Effects of Word N+1 Preview First-pass re-reading times were unaffected by word n+1 preview (F1(1,56)=1.53; p=.22; F2(1,40)=1.60; p=.21), however, all other durational measures showed a significant increase in inspection time following an invalid preview of word n+1 (first fixation duration: 264ms vs. 255ms: F1(1,56)=6.04; p<.05; F2(1,40)=4,93; p<.05; single fixation duration: 270ms vs. 258ms: F1(1,56)=11.10; p<.01; F2(1,40)=13.40; p<.01; last fixation duration: 263ms vs 253ms: F1(1,56)=10.77; p<.01; F21,40)=10.73; p<.01; gaze duration: 252ms vs 223ms: F1(1,56)=32.02; p<.001; F1(1,40)=42.26;

$p<.001$ and go-past time: 290ms vs. 254ms: $F1(1,56)=24.86$; $p<.001$; $F2(1,40)=26.14$; $p<.001$).

While first landing position on the word was unaffected by n+1 preview (both Fs<1), there was a reduced probability of skipping word n+1 if it had received an invalid preview prior to fixation (17% vs. 22%: $F1(1,56)=13.90$; $p<.001$; $F2(1,40)=14.66$; $p<.001$). There was some trend towards an interaction between word n+1 length and word n+1 preview ($F1(1,56)=2.99$; $p=.09$; $F2(1,40)=2.08$; $p=.15$); but identical previews were skipped more frequently than invalid previews irrespective of word length, although the effect was slightly stronger for 4-letter (34% vs. 26%: $F1(1,56)=11.33$; $p<.01$; $F1(1,40)=8.89$; $p<.01$) than for 6-letter words (11% vs. 8%: $F1(1,56)=4.18$; $p<.05$; $F2(1,40)=3.67$; $p=.06$). Like Experiment 2, this study provides another replication of word n+1 preview benefit (Rayner, 1975).

Effects of Word N+2 Preview The word n+2 preview manipulation failed to affect first landing position (both Fs<1), or inspection times on word n+1 (first fixation duration: both Fs<1.2; single fixation duration: $F1(1,56)=1.22$; $p=.27$; $F2<1$; last fixation duration: $F1(1,56)=2.03$; $p=.16$; $F2<1$; gaze duration: both Fs<1; go-past time: $F1(1,56)=1.52$; $p=.22$; $F2(1,40)=1.03$; $p=.32$; and first pass re-reading time: $F1(1,56)=1.78$; $p=.18$; $F2<1$). There was, however, evidence that an invalid preview of word n+2 increased the probability of skipping word n+1 by 3%. This trend approached significance by-subjects but was not typical of all items (18% vs. 21%: $F1(1,56)=3.84$; $p=.05$; $F2(1,40)=2.86$; $p=.09$). As with the first experiment in this chapter, this result provides evidence that something peculiar in the parafovea influenced targeting decisions made while fixating word n. But unlike the effect in Experiment 2, the invalid preview of word n+2 appears to have attracted attention directly towards it. This effect also differs from Experiment 2 in that it was not modulated by word n+1 availability – there was no interaction between the two preview manipulations ($F1(1,56)=2.42$; $p=.12$; $F2(1,40)=1.20$; $p=.28$) – suggesting that a double-attention shift is not a plausible explanation for it.

First fixation duration and single fixation duration both showed evidence of a three-way interaction between word n+1 length and the two preview manipulations, significant by-subjects but not by items in both cases (first fixation duration: $F1(1,56)=7.95$; $p<.01$; $F2(1,40)=2.12$; $p=.15$; single fixation duration: $F1(1,56)=5.19$; $p<.05$; $F2(1,40)=2.04$; $p=.19$). Follow-up analyses revealed that, while the two preview manipulations did not interact when word n+1 was a 6-letter word (first fixation duration: both $Fs<1.2$; single fixation duration: both $Fs<1$), they did when it was a 4-letter word, although again, these interactions were only significant by-subjects (first fixation duration: $F1(1,56)=10.51$; $p<.01$; $F2(1,40)=2.07$; $p=.15$; single fixation duration: $F1(1,56)=7.99$; $p<.01$; $F2(1,40)=1.54$; $p=.22$). Inspection of the means in Table 4.6 show that – for both measures – the difference between receiving an identical or invalid preview of word n+2 were small and unreliable if word n+1 had also received an invalid preview (first and single fixation durations, all

Fs<1.1), yet when word n+1 preview had been identical, an invalid preview of word n+2 increased both first and single fixation duration by 16ms and 14ms, respectively, both of which were significant by-subjects, but not by items (first fixation duration: $F1(1,56)=10.43$; $p<.01$; $F2(1,40)=1.87$; $p=.18$ and single fixation duration: $F1(1,56)=7.20$; $p<.01$; $F2(1,40)=1.40$; $p=.24$).

This pattern of results suggests – like Kliegl et al (2007) – that word n+2 preview influenced inspection times on word n+1 (albeit in a different measure). This effect was, however, clearly restricted to when word n+1 was a 4-letter word, to cases where it was parafoveally identical while fixating word n, and was apparently driven by only a subset of the items and therefore seems not to be a widely generalizable result.

No other interactions between any of the three variables were significant either by-subjects or by items (all ps>.10).

4.3.2.3. Word N+2

Fixation time and saccadic measures pertaining to word n+2 can be seen in Table 4.7.

Effects of Word N+1 Length There was a reverse word length spillover effect on word n+2, with longer inspection times following 4- compared to 6-letter words. While this effect was not seen in single or last fixation durations ($F1(1,56)=1.56$; $p=.22$; $F2(1,40)=2.01$; $p=.16$ and both Fs<1.2, respectively), it was clearly significant in first fixation duration (264ms vs.

256ms: F1(1,56)=5.74; p<.05; F2(1,40)=4.05; p<.05), gaze duration (298ms vs. 272ms: F1(1,56)=20.71; p<.001; F2(1,40)=25.01; p<.001) and go-past time (364ms vs. 328ms: F1(1,56)=16.46; p<.001; F2(1,40)=22.90; P<.001). There was also some tendency to regress less following a 6-letter word (56ms vs. 66ms: F1(1,56)=2.47; p=.12; F2(1,40)=2.85; p=.10).

Table 4.7 *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for Word N+2.*

	Four Letter Word				Six Letter Word			
	N+1 Identical		N+1 Invalid		N+1 Identical		N+1 Invalid	
	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid	N+2 Identical	N+2 Invalid
First Fix	269	260	257	270	257	254	258	253
Last Fix	259	250	249	260	259	253	260	256
Single Fix	268	262	256	269	262	258	260	254
Gaze	309	290	290	302	282	269	265	273
Go-Past	370	343	354	387	328	329	321	334
Re-Reading	61	53	64	86	45	59	56	61
Skip Prob	7	6	8	8	8	10	10	7
Landing	2.24	2.37	2.34	2.42	2.81	2.64	2.82	2.73

The higher incidence of skipping 4- compared to 6-letter words (see above) offers two potential explanations for these reverse effects. First, launch sites prior to skipping word n+1 will be more remote and offer less optimal conditions for word n+2 pre-processing compared with when word n+1 is fixated. Second, first fixation locations following word n+1 skipping may be less optimal since these fixations will follow longer saccades and therefore be more likely to – due to systematic range error – fall short of their intended targets (McConkie et al, 1988). Indeed, a leftward shift in first landing positions following 4- compared to 6-letter words was observed (2.34 vs. 2.75 character spaces: $F1(1,56)=19.82$; $p<.001$; $F2(1,40)=30.75$; $p<.001$).

Pollatsek, Juhasz, Reichle, Machacek and Rayner (2008) have also reported reverse word length spillover effects. However, despite implicating the above two mechanisms, Pollatsek et al suggest other mechanisms are likely to be jointly responsible. This conclusion was based on their finding that the spillover effect persisted when (a) the pre-target word (i.e., word n+1 in the present study) was fixated, and (b) the target word (i.e., word n+2 in the present study) landing position was optimal.

To account for these effects within the architecture of E-Z Reader, Pollatsek et al propose that integration failure will occur more frequently following 4- rather than 6-letter words. They suggest that this is plausible given that 4-letter words will have more neighbours and therefore will be more likely to be misidentified than 6-letter words. It remains possible, however, that these effects could also be explained from a parallel perspective if one assumes that longer fixation durations on word n+1 afford more time for parafoveal processing of the upcoming word; unfortunately SWIFT has not been used to simulate such effects and so the validity of this reasoning has not been directly tested. Such an assumption is however, within the theoretical makeup of a parallel model.

Effects of Word N+1 Preview A trend in first-pass re-reading time suggests that more time was spent re-reading earlier text if word n+1 had received an invalid preview (67ms vs. 55ms: $F1(1,56)=3.30$; $p=.07$; $F2(1,40)=3.78$; $p=.06$). No other measure demonstrated a continuing cost of having received an invalid preview of word n+1 (gaze duration: $F1(1,56)=1.08$; $p=.30$; $F1(1,40)=1.47$; $p=.23$; all other measures: all $Fs<1$). This trend in re-reading time is qualitatively similar to the effects observed in Experiment 2, with the longer duration associated with a previously invalid preview, although it is clearly a weaker effect. Nevertheless, for the reason outlined previously, such a result is not consistent with E-Z Reader unless it is assumed that it was caused by an erroneous fixation of word n+2 followed by quick error-correcting regression, but only in those cases when word n+1 had received an invalid preview. Since this orthographic spillover effect is smaller in this experiment when word n+1 was longer (and less likely to be skipped) than in Experiment 2 where word n+1 was always 3-letters in length, it is consistent with this hypothesis. However, such an explanation should also predict an interaction between word n+1 length and preview in the present experiment, and this was clearly absent (both $Fs<1.2$).

Effects of Word N+2 Preview There were no main effects of n+2 preview on word n+2 (single fixation duration: F1<1; F2(1,40)=2.17; p=.15; first-pass re-reading time: F1<1; F2(1,40)=2.17; p=.15; all other Fs<1). While there was some indication that word n+1 length might have modulated an effect of n+2 preview in first landing position, the interaction was clearly nonsignificant by-subjects (F1(1,56)=2.58; p=.11; F2(1,40)=5.25; p<.05). Unsurprisingly, therefore, neither the earlier first landing positions following the 6-letter words nor the later first landing positions following the 4-letter words achieved significance (F1(1,56)=1.40; p=.24; F2(1,40)=2.32; p=.13 and F1(1,56)=1.65; p=.20; F2(1,40)=1.33; p=.25, respectively).

There was a trend towards an interaction between the two preview manipulations in first fixation duration (F1(1,56)=3.15; p=.08; F2(1,40)=3.03; p=.09) and this was significant in gaze duration (F1(1,56)=6.28; p<.05; F2(1,40)=5.64; p<.05) and marginally significant in go-past time (F1(1,56)=3.60; p=.06; F2(1,40)=5.52; p<.05). The nature of these interactions can be seen below in Figure 4.10.

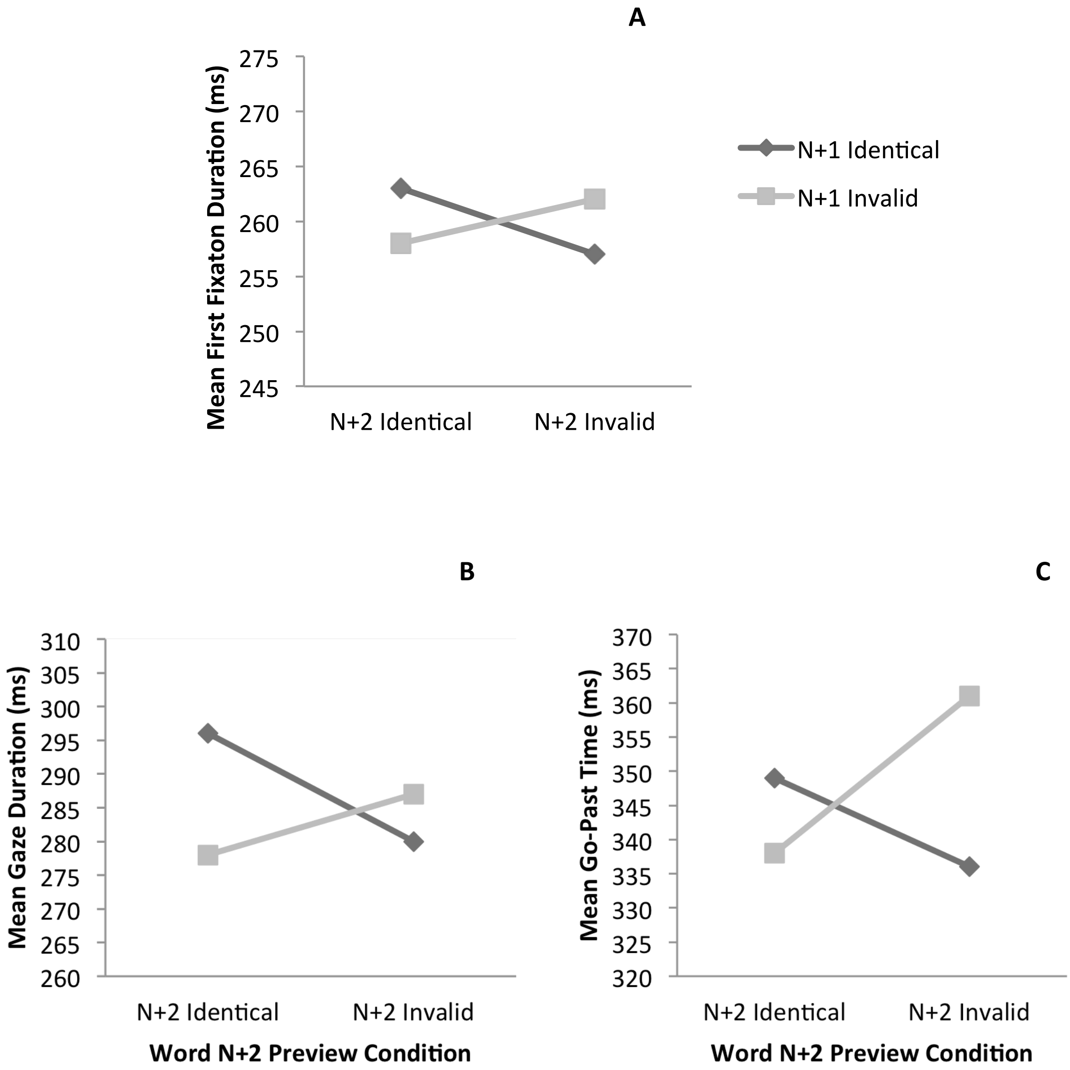


Figure 4.10. *Mean (A) First Fixation Duration, (B) Gaze Duration and (C) Go-Past Time (ms) on Word N+2 for the Two Word N+1 Preview Conditions as a Function of Word N+2 Preview.*

As can be seen, the availability of word n+1 modulated word n+2 preview benefit. An invalid preview of word n+2 seems to have encouraged shorter inspection times on word n+2 if word n+1 had received a valid preview, but longer inspection times if word n+1 had also received an invalid preview. It was clear that, for first fixation duration at least, these modulations were nonsignificant (n+1 identical: $F1(1,56)=2.60$; $p=.13$; $F2(1,40)=2.30$; $P=.13$, n+1 invalid: both $Fs<1$). However, while the 9ms increase in gaze duration was not significant when word n+1 had received an invalid preview ($F1(1,56)=2.01$; $p=.16$; $F2(1,40)=1.31$; $p=.26$), the 23ms increase in go-past time was significant by-subjects and marginally significant by-items ($F1(1,56)=4.05$; $p<.05$; $F2(1,40)=3.53$; $p=.06$), suggesting that if both words n+1 and n+2 were parafoveally invalid while fixating word n, readers tended to spend more time inspecting both n+2 and earlier parts of the text prior to moving forward. If, however, word n+1 had been parafoveally available while fixating word n, an invalid preview of word n+2 significantly reduced gaze duration ($F1(1,56)=6.69$; $p<.05$; $F2(1,40)=4.91$; $p<.05$), although the corresponding decrease was not significant in go-past time ($F1(1,56)=1.04$; $p=.31$; $F2(1,40)=1.35$; $p=.26$).

These results are important since they demonstrate the existence of word n+2 preview benefit on word n+2 when (a) the intervening word n+1 contained more than 3-letters, and (b) word n+1 had been parafoveally unavailable while fixating word n, which – according to the E-Z Reader model – should have prohibited the movement of attention from word n+1 to word n+2. Word n+2 preview influencing fixation durations on word n+2, strongly implicates parallel processing while reading.

There was no further evidence of any interactions between any of the three variables (all $ps>.09$).

4.3.2.4. Spillover Region

Effects of Word N+1 Length As can be seen in Table 4.8, any delayed effects of word n+1 length had dissipated by the time the spillover region was reached (first pass re-reading time: F1(1,56)=2.61; p=.11; F2(1,40)=1.12; p=.30; all other measures: Fs<1).

Table 4.8. *Fixation Time Measures (ms) and First Landing Positions (character spaces) for the Spillover Region.*

	Four Letter Word				**Six Letter Word**			
	N+1 Identical		**N+1 Invalid**		**N+1 Identical**		**N+1 Invalid**	
	N+2 Identical	**N+2 Invalid**	**N+2 Identical**	**N+2 Invalid**	**N+2 Identical**	**N+2 Invalid**	**N+2 Identical**	**N+2 Invalid**
First Fix	243	251	238	242	245	249	237	240
Gaze	462	466	469	473	469	472	467	481
Go-Past	515	553	533	523	512	541	527	522
Re-Reading	53	87	65	50	43	70	60	40
Landing	4.18	3.75	4.15	4.24	4.27	4.01	4.09	3.96

Effects of Word N+1 Preview The inflated inspection times on word n+2 following an invalid preview appear to have permitted more time to extract information from the spillover region, resulting in shorter first fixation durations in that zone (239ms vs. 247ms: F1(1,56)=5.79; p<.05; F2(1,40)=7.79;

$p<.01$). This effect was confined to first fixation duration and was not evident in any other measure (all $Fs<1.1$)

Effects of Word N+2 Preview While there was no main effect of word n+2 preview in any of the durational measures (first fixation duration: $F1(1,56)=1.95$; $p=.16$; $F2(1,40)=2.83$; $p=.10$; gaze duration: both $Fs<1$; go-past time: $F1(1,56)=2.37$;$p=.13$;$F2(1,40)=1.48$; $p=.23$; first pass-re-reading time: both $Fs<1$), there was a tendency for first fixations to land earlier in the spillover region if word n+2 had previously received an invalid preview (3.99 vs. 4.17 character spaces: $F1(1,56)=4.63$; $p<.05$; $F2(1,40)=3.58$; $p=.06$). As can be seen in Figure 4.11, a numerical trend towards an interaction between the two preview manipulations, suggests that this effect was strongest when word n+1 had received an identical preview ($F1(1,56)=5.22$; $p<.05$; $F2(1,40)=7.13$; $p<.05$) rather than an invalid one (both $Fs<1$); although the qualifying interaction was clearly nonsignificant by-subjects ($F1(1,56)=1.75$; $p=.19$; $F2(1,40)=3.93$; $p=.05$). Since word n+2 preview did not influence word n+2 skipping rates, these results suggest that the word n+2 word change had been identified and a more cautious reading strategy was adopted immediately thereafter, especially when word n+1 had been available.

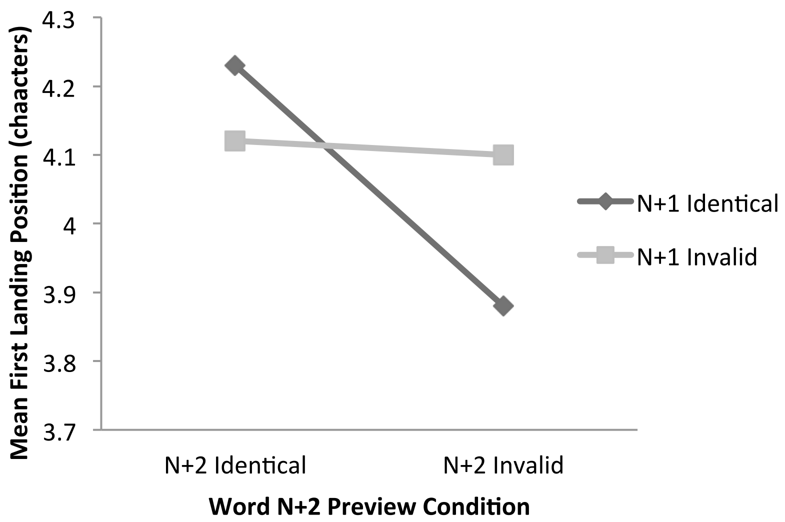


Figure 4.11. *Mean First Landing Position (characters) in the Spillover Region for the Two Word N+1 Preview Conditions as a Function of Word N+2 Preview.*

Finally, as can be seen in Figure 4.12, there was an interaction between the two preview manipulations in go-past time (F1(1,56)=5.79; p<.05; F2(1,40)=6.14; p<.05), apparently driven by a similar interaction in re-reading time (F1(1,56)=7.73; p<.01; F2(1,40)=8.68; p<.01). Pairwise comparisons revealed that while there was no significant effect of word n+2 preview when word n+1 had received an invalid preview (go-past time: both Fs<1; re-reading time: F1(1,56)=2.94; p=.09; F2=(1,40)=1.56; p=.22,), the increases following an invalid preview of word n+2 when word n+1 was identical were significant, both in go-past time ((F1(1,56)=5.69; p<.05; F2(1,40)=4.48; p<.05) and first-pass re-reading time (F1(1,56)=5.52; p<.05; F2(1,40)=5.90; p<.05).

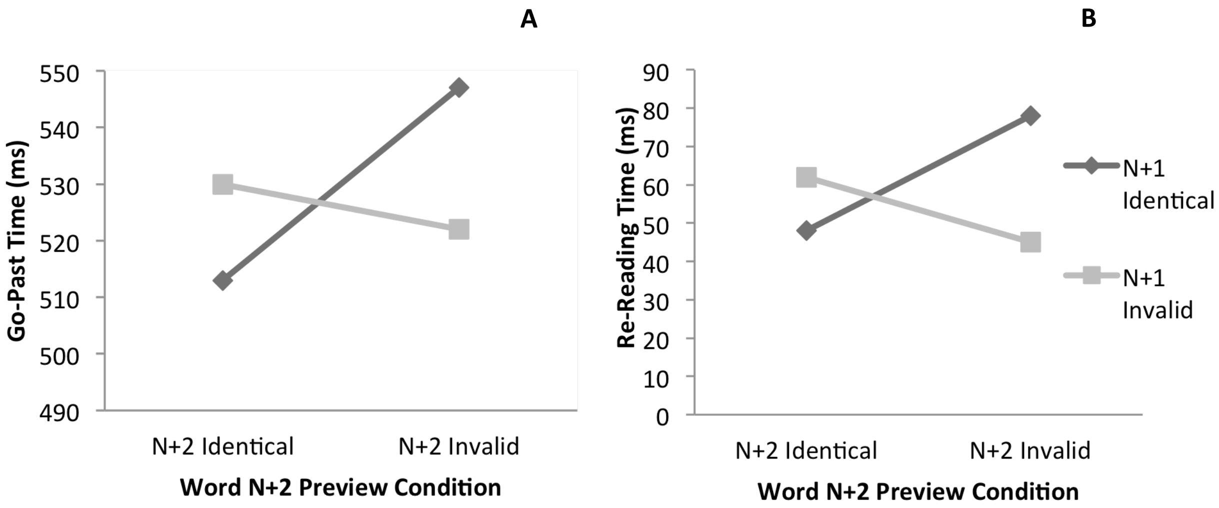


Figure 4.12. *Mean (A) Go-Past Time and (B) First-Pass Re-Reading Time (ms) in the Spillover Region for Identical and Invalid Previews of Word N+1 as a Function of Word N+2 Preview.*

Again, these interactions are the mirror opposite of those present in first fixation duration, gaze duration and go-past time on word n+2, suggesting a trade-off existed between the time spent inspecting word n+2 and time spent in the spillover region.

4.3.3. General Discussion of Experiment 3

This study sought to test the range over which word n+2 preview benefit might occur. Previous studies using longer words (greater than 3 characters) in position n+1 have either failed to provide a strong test as a consequence of the use of low frequency words (Rayner et al, 2007), or have included word n+1 lengths that, due to acuity constraints, render word n+2 pre-processing far less likely (e.g., up to 10-letters in Angele et al, 2008).

As discussed above, evidence pertaining to word n+2 preview benefit over a longer range should help to distinguish the predictions from the two most prominent models of eye movement control during reading. While SWIFT should continue to predict word n+2 preview benefit over a longer range (albeit to a lesser extent), the mechanism through which word n+2 preview benefit could be explained in the E-Z Reader model should typically inhibit such effects occurring over a longer range. This mechanism – a double attention shift – is contingent upon word n+1 being fully processed (allowing attention to then move onto word n+2) in the lag between word n identification and a saccade out of it being executed. Consequently, the E-Z Reader model should only predict word n+2 preview benefit with an 'easy' to process parafoveal word n+1.

Despite the increased length of word n+1 in the present study, several effects of word n+2 preview were observed. These effects were not localized to word n+2. Rather, trends and/or effects were also apparent on the two preceding words and in the spillover region.

Three word n+2 parafoveal-on-foveal effects were observed, all of which suggest that something peculiar in the position of word n+2 attracted attention. First, there was evidence that when word n+1 received an identical preview, something peculiar in the position of word n+2 appears to have inhibited regressions from word n, with less re-reading time in these cases. Second, when word n+1 received an invalid preview, there was a tendency for an invalid preview of word n+2 to reduce last fixation duration on word n. Third, like Angele et al (Exp 1, 2011) and Radach et al (2013), there was a tendency for higher skipping probabilities of word n+1 when word n+2 had received an invalid preview. Taken together, these results appear to exhibit the sort of attraction mechanism previously suggested by Kennedy (1998) and Hyönä and Bertram (2004), preventing regressions, shortening last fixation duration and encouraging the skipping of an intervening word. Importantly, the last two results were found when word n+1 had received an invalid preview, indicating that they could not have been caused by an attention shift from word n+1 to word n+2 following successful lexical activation of word n+1.

While the present word n+1 skipping effect lies in the opposite direction to that found in Experiment 2, it should be recalled that the origins of that effect appeared to have been seated in an increased probability of skipping word n in the presence of word n+2 illegality. This consequently resulted in an increased word n+1 fixation probability. Thus both effects appear to suggest that something peculiar in the parafovea, at some point prior to passing word n, initiated an attraction mechanism.

With respect to the parafoveal-on-foveal effects shown in durational measures, it will be recalled from Chapter 2 that these effects lie in the opposite direction to those reported by Kliegl et al (2007). They are, however, consistent with the effects present in Experiment 2. As previously discussed, variation in foveal word length appears to often affect the mode of expression of parafoveal-on-foveal effects (e.g., Hyönä & Bertram, 2004; Kennedy & Pynte, 2005; Kennedy et al, 2002) and variations between languages (e.g., Kennedy & Pynte, 2005; Risse & Kliegl, 2011) has also been implicated as a possible contributing factor. While the present study does not adjudicate between these possibilities, it does provide further evidence for the presence of long range parafoveal-on-foveal effects, with effects from word n+2 observed even when word n+1 was four or six characters in length, and with generally similar effects across these lengths.

Of course, proponents of the serial perspective might argue that since this collection of parafoveal-on-foveal trends and effects are orthographic in nature, they could have been driven by the low level attentional scan detecting upcoming irregularity. Indeed, this was exactly the argument Angele and Rayner (2011) used when they found evidence that word n+2 preview influenced word n+1 fixation probability. It is difficult to assess whether such a mechanism could be responsible for these effects without it being implemented in the E-Z Reader model and without detailed simulations. While the possibility cannot be discounted that a low level scan could be driving these effects, such a mechanism, initiating a skip of an unidentified word, does seem somewhat incongruent in the context of a model that assumes a serial sequential process of word recognition as the engine driving the eyes through text.

As discussed in the Introduction, the presence of word n+2 effects over a longer range might help to determine whether so-called delayed parafoveal-on-foveal effects can be accounted for by a failed skip of word n+1 followed by a stay and process response. Unfortunately, the present set of results does not help adjudicate this question. While there was some evidence for longer inspection times on word n+1 if word n+2 had received an invalid preview, these results were restricted to instances in which word n+1 was an available 4-letter word, and were clearly restricted to a subset of the items. The skipping probability for word n+1 in these cases – approximately one-third – is sufficiently high for it to remain conceivable that delayed parafoveal-on-foveal effects are driven by a proportion of skips that happen to fall short of their intended word n+2 target. The finding that the effect was restricted to cases where word n+1 was available also appears to lend some support to such a hypothesis. This explanation for delayed parafoveal-on-foveal effects must therefore remain a possibility.

Unlike Experiment 2 and the majority of research into word n+2 preview effects to date (e.g., Angele et al, 2008; Angele & Rayner, 2011; Exp1; Kliegl et al, 2007; Rayner et al, 2007), the present study obtained word n+2 preview effects on word n+2. The direction of these effects depended upon whether word n+1 had received a valid or invalid preview prior to exiting word n. When word n+1 had received an identical preview, an invalid preview of word n+2 reduced gaze duration on word n+2 compared to when it had received an identical preview. This result is counterintuitive, since denying word n+1 an identical preview prior to fixation typically results in longer, not shorter, durations once that word is eventually fixated (e.g. Rayner, 1975). It is important to remember, however, that unlike word n+1 preview benefit, in this experiment, an intervening word (word n+1) can be fixated prior to fixating word n+2, and during the fixation on word n+1, word n+2 has an identical preview and so can accrue preview benefit. A more complex pattern of effects might therefore be expected from the manipulation of word n+2 preview, since it will be modulated by any inspection differences on word n+1. And indeed, inspection times on word n+1 were inflated when the word to the right had received an invalid preview, and this could have contributed towards the effect. Specifically, it could be suggested that there was a tradeoff between word n+1 processing time (inflated as a consequence of an invalid preview of word n+2) and word n+2 processing time (now reduced as a consequence of the extended time spent fixating word n+1). However, while such a mechanism might contribute to this effect, it is unlikely to be the sole factor, since word n+2 effects on word n+1 were restricted to 4-letter words while the effect on word n+2 was present for both lengths, although it was, numerically stronger following a 4-letter word n+1.

While an explanation for this pattern of results is elusive at present, it is interesting to note that, again following an identical preview of word n+1, an invalid preview of word n+2 appears to have resulted in earlier first landing positions within the spillover region, suggesting a more cautious reading strategy had been adopted. It is clear therefore, that when word n+1 received an identical preview, word n+2 preview influenced reading strategy.

A classical word n+2 preview benefit was also observed in go-past time on word n+2 when word n+1 had also received an invalid preview. This effect suggests that readers spent more time refixating and regressing when both critical words had received an invalid preview compared with when only word n+1 had received an invalid preview. This effect is particularly interesting since it occurred when the condition of a double attention shift could not be satisfied owing to the simultaneous masking of both words.

This last result, together with the strong trend for parafoveal-on-foveal effect in last fixation duration on word n when word n+1 had received an invalid preview, indicate that a double attention shift is unlikely to be driving these effects. Since this is the primary mechanism to which word n+2 preview effects can be attributed in the E-Z Reader model, the present set of results do not seem compatible with that model. Of course, these effects were orthographic in nature, allowing the proponents of E-Z Reader to suggest that they might arise from a low level attentional scan detecting the upcoming irregularity. However, as discussed above, whether or not such a mechanism could account for the present patterns of results can only be determined once that mechanism is included within an implementation of the E-Z Reader model.

The present sets of results represent the first time a word n+2 preview benefit has been obtained on word n+2 (or thereafter) when word n+1 contained more than 3-letters. It was therefore a surprise to find word n+2 preview effects on word n+2 in Experiment 3 but not in 2. The reason for this difference is not particularly clear, but it might be suggested that attention is distributed differently when a potential word skip is involved. It could also be related to the fact that short function words were not generally laden with semantics, so there is less imperative to process word n+2 compared to the situation with an adjective, where typically the interpretation requires a combination of both words. While the present study does not adjudicate between these (or other) possibilities, it has demonstrated the importance of not restricting word n+1 to a specific word class, or length.

An unexpected result of the present study was the finding that the length of the word n+1 ‘mask’ appears to modulate the parafoveal-on-foveal effects expressed on word n. This is a significant finding, as it suggests that variation in the length of parafoveal irregularity might regulate the expression of parafoveal-on-foveal effects – with some appearing in an orthodox, while others appear in an unorthodox direction – and such variations might act to cancel out orthographic parafoveal-on-foveal effects entirely, explaining why some researchers fail to obtain the effects. Since the qualifying interaction for this effect between word n+1 length and word n+1 preview was significant by-subjects but only marginal by-items, replication is obviously desirable. It should be noted, however, that this is not the first study to find evidence that parafoveal word length might modulate the expression of parafoveal-on-foveal effects (e.g. Hyönä & Bertram, 2004; Kennedy et al, 2002), therefore, previous mistrust in parafoveal-on-foveal effects due to their inconsistent expression (e.g., Rayner et al, 2003) may have been premature. Future research should clearly take care to ensure that such potential confounds are avoided when investigating parafoveal-on-foveal effects.

Similarly to experiment 2, a trend towards a word n+1 preview spillover effect was present on word n+2. As previously discussed, this pattern of data is incongruent with a model like E-Z Reader, that assumes the decision to plan a saccade to the next word in text is coupled with completion of the orthographic and phonological stages of word recognition (L1). If this were the case then the denial of orthographic preview benefit should only affect the word undergoing the manipulation and never spillover. Proponents of the serial perspective might suggest that these effects reflect an accidental overshoot of word n+1 followed by a regression on the occasions that the input had changed from an invalid to an identical preview. There was, however, a lack of empirical support for such an explanation, since following this logic, it should also be predicted that an interaction would be present between word n+1 length and word n+1 preview. Specifically, it has been suggested that overshoots are most prevalent following a short saccade (i.e., saccades that are less than the preferred saccadic length of 7 characters;

McConkie et al, 1988), therefore, there should be fewer occasions in which a saccade will overshoot a 6- compared to a 4-letter word. No such interaction was observed in the present set of results.

Theoretically, the above explanation is also inconsistent with the assumption that a mislocated fixation is followed by a stay and process response, as suggested by proponents of the serial perspective (e.g., Drieghe et al, 2008). For the above explanation of word n+1 spillover effects to work, the mislocated fixation must be followed by a corrective saccade back to the intended target word, and not a stay and process response. Potential responses to mislocated fixations is a topic that will be returned to in Chapter 6, but for now, it seems unlikely that mislocated fixations can account for the present set of word n+1 spillover effects. While it is acknowledge that the spillover effect in the present set of results was only marginally significant, the same effect was present and achieved significance in go-past time in Experiment 2. It will be interesting to see whether these trends and effects can be replicated, since if they can, such results clearly present a challenge for the E-Z Reader model of eye movement control.

Finally, this experiment provided a replication of the reverse word length spillover effect, with increased inspection times associated with short preceding, compared with long preceding, words. As discussed above, Pollatsek et al (2008) were able to simulate these effects using the E-Z Reader model by including the auxiliary assumption that short words will be misidentified more frequently than long words, and this is often only noticed once the subsequent word is processed, at which point the error is detected and inspections increase accordingly. While this theory cannot be discounted, a more parsimonious explanation seems to be that longer durations on word n+1 permit more pre-processing of word n+2, resulting in a trade-off in processing time between words n+1 and n+2. While not yet simulated using the SWIFT model, such an explanation does appear to be most compatible with a model that allows multiple words to be processed simultaneously, and which allows this processing to influence upcoming targeting decisions.

Again, as a result of the more specific predictions derived from it, this discussion has centered on whether the E-Z Reader model is capable of accounting for the variety of effects displayed in this experiment. This model is the most advanced and extensively-tested model of the serial variety and provides the perfect platform on which we can assess whether the present set of (apparently parallel) effects can be accounted for with a serial framework. Very little, however, has been said regarding whether its main competitor – the SWIFT model – would also do a good job of simulating these effects. While the present set of effects appear to fit most parsimoniously within a framework that permits the lexical processing of multiple words simultaneously, the present version of SWIFT (SWIFT 3, Schad & Engbert, 2012) would undoubtedly fail to capture the complexity of the present results. While recent simulations show that SWIFT is capable of simulating orthodox word n+2 preview effects on words n+1 and n+2 (i.e., longer durations on those words when word n+2 received an invalid preview while fixating word n; Risse et al, 2014), these results obviously do not align with the finding of shorter durations on word n+2 following an invalid preview, if word n+1 had been visible while fixating word n. Equally, while it is argued that SWIFT should be able to reproduce weak word n+2 parafoveal-on-foveal effects (e.g. Risse et al, 2008), the model does not predict the effects of 'attraction' seen here. It is therefore clear that both models would fail to simulate the complex pattern of effects observed in the present study. But despite SWIFT apparently failing to capture the complexities of these results, the overall patterns appear more in harmony with a model that assumes multiple words can be lexically processed in a parallel fashion.

As suggested above, obtaining word n+2 preview effects over a longer range greatly reduces the probability that they can be accounted for by a double attention shift. However, in a recently published paper, Schotter et al (2014) presented a series of simulations demonstrating that the E-Z Reader model is capable of simulating a modest word n+2 preview benefit on word n+2 and, importantly, the time previewing word n+2 was unrelated to word n+1 length and frequency. In fact, they report that, when a 5-letter word was entered as word n+1 into the E-Z Reader model simulation, word n+2 pre-processing occurred on 19% of occasions, with a mean word n+2 pre-processing duration of 141ms (this time includes 25ms saccade execution time and 50ms eye-mind lag; Schotter et al, 2014). In their simulations, however, word n+1 was always available for inspection, since this is a necessary prerequisite for attention to move to word n+2 within the model. These simulations cannot, therefore, account for the effects seen here when word n+1 had received an invalid preview. Indeed, while these simulations suggest that orthodox word n+2 preview effects can arise on word n+2 within the architecture of the E-Z Reader model, the simulations fail to capture the complexity of the present set of results, such as the unorthodox parafoveal-on-foveal effects or the unorthodox word n+2 effect following a valid preview of word n+1.

4.4. Conclusion

The results from these experiments show evidence of several word n+2 preview effects spanning words n, n+1, n+2 and the spillover region. Since the effects were expressed in inspection patterns related to word n+2 in the present experiment but not in Experiment 2, this suggests that word n+1 length might be critical in modulating effects present on word n+2. The pattern across the two experiments was complex, potentially driven in part by interactions between processing times on each of the critical words. For this reason it seems unlikely that either the E-Z Reader or the SWIFT models of eye movement control will be capable of simulating the totality of this pattern. The effects reported here do, however, seem to fit most parsimoniously with a model that assumes lexical processing is distributed across multiple words simultaneously.

CHAPTER 5

Plausibility Preview Effects from Words N+1 and N+2

5.1. Introduction

The boundary paradigm has been used extensively over the past forty years to investigate the extent and nature of parafoveal processing during reading. This research has shown that fixation durations are reliably shorter following orthographically and phonologically related previews compared to unrelated previews, suggesting that both of these word features can be activated prior to direct fixation (e.g., Rayner, 1975 and Pollatsek, Lesch, Morris & Rayner, 1992, respectively; see Chapter 2). Correspondingly, the failure of early studies to provide evidence for a semantic preview benefit – that is, where previewing a semantically related word aids subsequent identification – has been taken as evidence against the routine extraction of a parafoveal word's semantic code (e.g., Altarriba, Kambe, Pollatsek & Rayner, 2001; Hyönä & Häikiö, 2005; Rayner, Balota & Pollatsek, 1986). Such findings have been considered theoretically important since they align with the premise that typically only the early stages of word recognition have time to occur on a parafoveal word prior to its direct fixation (i.e., orthographic and phonological code extraction), with the later stage of semantic processing typically being associated with direct foveal inspection. However, full parafoveal identification must occur on occasions, as this accounts for some (but not all) skipping behaviour during reading (e.g. Drieghe, Rayner & Pollatsek, 2005). But for cases where the parafoveal word is fixated, the apparent absence of semantic preview benefit has been taken as evidence that fixation location and lexical identification are tightly coupled during reading (e.g., Rayner et al, 1986; Rayner, White, Kambe, Miller & Liversedge, 2003).

The absence of semantic preview benefit is consistent with the E-Z Reader model (Reichle, Warren & McConnell, 2009). As will be recalled from Chapter 1, in this model, a saccade is programmed to the next word in text when the currently fixated word achieves the L1 stage of word recognition. Thus the amount of time available for parafoveal pre-processing is necessarily restricted to the time required to plan and execute a saccade to the next word in text minus the time required to complete lexical processing on the fixated word (i.e., reach stage L2). Under these tight time constraints, typically only the very early stages of word recognition should have time to complete on a parafoveal word prior to direct fixation. Furthermore, according to the model, if time does permit the extraction of the semantic code, this will not always be reflected as a semantic preview benefit on the target word, since a parafoveally-identified word will be skipped if identified early enough (i.e., before the saccade is committed to action; see Chapters 1 and 2 for more detail). Thus, within the model, not only do the tight time constraints typically prevent the parafoveal extraction of a word's semantic code, but when it is extracted, the mechanism which produces word skipping will typically prevent semantic preview benefit from being expressed on the target word.

In contrast, semantic preview benefit appears to fall naturally within the theoretical remit of a processing gradient model (e.g., SWIFT; Schad & Engbert, 2012), with multiple words processed simultaneously, potentially up to the stage where semantic information could be extracted. With such a clear divide in the nature of the predictions derived from these two classes of model, it is perhaps unsurprising that researchers have directed their attention towards the question of whether semantic preview benefit exists.

As will be recalled from the Chapter 2, the first study to investigate semantic preview benefit was conducted by Rayner et al (1986). They manipulated the parafoveal preview of a target word such that prior to fixation it received either an identical preview, a semantically related or unrelated preview, or a visually similar nonword preview. Despite obtaining evidence for facilitation following the visually similar nonword preview, there was no evidence to suggest that a semantically related preview reduced target word inspection times compared to an unrelated preview, and therefore no evidence for a semantic preview benefit. Since Rayner et al (1986) were able to demonstrate that the semantic associates used in their task produced facilitation in a pronunciation task, they concluded that parafoveal processing does not typically extend to the semantic level during reading. This finding appears to be reliable since Rayner, Schotter & Drieghe (2014) have recently replicated Rayner et al's (1986) results showing an absence of semantic preview benefit using an almost identical item set.

It will also be recalled from Chapter 2, however, that Rayner et al's conclusions rest on one important assumption: that semantic associate previews in boundary paradigm experiments facilitate target word processing in the same way that semantic associate primes have been shown to facilitate responses to target words in isolated word tasks such as lexical decision (e.g., Meyer & Schvaneveldt, 1971) or pronunciation tasks (e.g., Rayner et al, 1986).

There are good reasons, however, to question whether these two types of task should be considered to be equivalent. First, in tasks such as lexical decision, the target word does not change part way through processing. Typically one word is presented first, and this word has the potential to facilitate the lexical processing of a subsequently presented word. In a gaze contingent display change paradigm, however, there *is* a change in the nature of the word 'object', with it initially adopting one identity and this being replaced by another when directly fixated. Since this change will frequently occur with the word having undergone some level of processing, it is possible that the change might produce some form of interference (e.g. Murray, Rayner & Wakeford, 2013; Risse & Kliegl, 2013).

Second, there is a change in meaning that accompanies a change from one semantically related preview to another that might also add to any interference. Indeed, it should not be assumed that semantically related words carry the same meaning; they are merely associated on some level. It is possible, therefore, that a word change from one semantic associate to another might disrupt on-going sentence interpretation, especially when there is little correspondence in meaning between the preview and target word. For example, arms-legs, north-south and rattle-bottle (taken from Rayner et al, 1986) are all considered to be semantic associates, and therefore, according to Rayner et al (1986), should facilitate target word processing. However, as these examples demonstrate, although semantically associated, all carry very different (often opposite) meanings, and this too could be responsible for disrupting the reading process.

Finally, word changes in semantic preview benefit experiments typically occur mid-sentence, that is, after some level of context has been established. This means that if the semantic code of a parafoveal word has been extracted prior to fixation, it might also have been integrated at the discourse level. Thus, even a relatively close correspondence in meaning between two semantically associated words might result in some interference, with the extent of this dependent upon how much the prior sentence context might have been appropriate for, more predictive of, or more consistent with one meaning rather than the other.

Thus, a change in meaning between the two semantically associated words could interfere with, rather than facilitate target word processing. Furthermore, the origins of this interference might be present at various different levels of processing – lexical identity, lexical meaning and sentence or discourse meaning. It is possible, therefore, that meaning based interference effects may cancel out any semantic preview benefit that might have accrued on a word prior to fixation.

Several researchers have explored the possibility of meaning-based interference effects within semantic preview benefit experiments. As discussed in Chapter 2, Rayner et al (1986) reanalysed their data including only those sentences rated as most similar in meaning; no evidence for a semantic preview benefit was observed on this reduced dataset. However, not only will this step have reduced power in the study, a rating of *overall* sentence meaning will not necessarily capture the disruption caused to on-going sentence interpretation at the point at which the meaning change was first encountered. Thus, interference effects may still have been present at the critical point in the sentence.

Altarriba et al (2001) also attempted to avoid potential meaning-based interference effects by using previews that were direct translations of target words with fluent Spanish-English bilinguals; again, no evidence for a semantic preview benefit was observed. It is, however, possible that language-switching costs (Meuter & Allport, 1999) may have negated any semantic preview benefit in this study. Also, there is the question of whether semantic facilitation spreads between the lexica of two different languages.

Some support for an influence of meaning based interference effects in semantic preview benefit experiments was recently reported by Hohenstein, Laubrock & Kliegl (2010). For normally presented, lower case words, Hohenstein et al report obtaining a semantic preview benefit when preview availability was restricted to the first 125ms of an initial fixation on the pre-target word; this effect was absent at the shorter durations of both 35ms and 80ms. When the semantically related and unrelated previews were presented in bold to increase the saliency of the preview, there was again evidence for a semantic preview benefit, but this time only when the preview was available for the first 80ms of the initial fixation on the pre-target word, not at the longer duration of 125ms. A possible interpretation of these results is that some minimal amount of time is required before the semantic associate can become activated. However, if available for too long, as was perhaps the case in the condition of increased saliency, the meaning of the semantic associate might become available in a way that begins to interfere with processing of the target word upon fixation. Unfortunately, the experimental conditions of this study necessitated the word change occurring while the eye was stationary, so the change did not occur during saccadic suppression, and a word change under such conditions might have affected how attention is typically distributed during reading.

This study was the first, however, to provide some evidence that meaning based interference might play a role in semantic preview benefit experiments. Thus, while it is clear that researchers have been conscious of the potential confound of meaning based interference in semantic preview benefit experiments, testing this possibility has proved difficult.

Two recently-reported studies do, however, lend some support to the hypothesis that semantic preview benefit can be found when meaning correspondence between preview and target is sufficiently high. First, Hohenstein et al (2013) found a semantic preview benefit for sentences like: "*With the excavation, bones came to light"*, where the semantic associate "skulls" turned into "bones". Within the context of the sentence, "bones" and "skulls" carry a similar meaning, reducing the potential for meaning based interference effects. Since Hohenstein et al did not provide a full item list, it is unknown whether all semantic associates in their experiment shared such a meaning correspondence, but if they did, this could certainly explain why they obtained evidence of semantic preview benefit while others have failed: their semantic associates might have produced less meaning based interference.

Perhaps the most convincing evidence to date that meaning based inference might cancel out the expression of a semantic preview benefit was reported by Schotter (2013). As outlined in Chapter 2, Schotter obtained semantic preview benefit with synonymous word pairs (Experiments 1 and 2), but not with semantically associated word pairs (Experiment 2). Furthermore, post hoc analyses revealed a negative correlation between similarity of meaning and inspection time on the target word. The closer the correspondence in meaning between the two related words, the faster the target word was subsequently processed. Schotter's results are therefore consistent with the general premise that meaning based interference might be responsible for failed attempts to uncover a semantic preview benefit.

The experiments presented in this Chapter also investigate the question of whether there is higher level pre-processing but in a novel way. Rather than searching for semantic preview benefit *per se*, Experiments 4 to 6 tested the presence of semantic processing by looking for differing levels of semantic *interference*. This was achieved by using the gaze contingent display change paradigm to manipulate parafoveal plausibility, and measuring whether variations in plausibility influence the subsequent inspection time of target words. If differences do arise, then this suggests that higher-level pre-processing has occurred at an advanced level prior to direct fixation, and this would potentially implicate a much looser coupling between fixation location and the locus of attention than can currently be accommodated within a serial architecture. Experiment 4 tested for plausibility-related preview effects originating from word n+1 and Experiments 5 and 6 tested the range over which such effects might arise by testing for the presence of these effects in the position of word n+2.

5.2. EXPERIMENT 4

Manipulating parafoveal plausibility in order to determine when the meaning of a word becomes available is not new. In a series of same-different matching studies conducted by Murray and colleagues (Kennedy, Murray & Boissiere, 2004; Murray, 1998; Murray, 2006; Murray & Rowan, 1998), participants were required to read sentences and to press a button, triggering another sentence to be displayed; the task being to indicate whether the two sentences were the same or different. These studies showed, during reading of the initial sentence, effects of the plausibility of the combination of the initial noun phrase with the verb, for example, "*The hunters stacked….*" vs "*The bishops stacked…*", and in a number of the studies, this was reflected not only in fixations falling on the verb, but also in some eye movement measures before the verb was directly fixated, suggesting the extraction of meaning from words in the parafovea. However, Rayner et al (2003) report being unable to replicate one of Murray et al's findings in a reading study and suggest that these results might have been task specific.

Starr and Inhoff (2004; Experiment 1) also investigated the consequences of providing a contextually inappropriate word to the right of fixation. They initially presented a critical word as either itself, a contextually inappropriate word, or a legal or illegal nonword. In addition to finding clear orthographic parafoveal-on-foveal effects, a trend also emerged in which a contextually inconsistent word in the parafovea reduced gaze duration by 22ms on the pre-target word compared to an accurate preview. However, a subsequent analysis excluding the 45% of cases where fixations fell near the end of the pre-target word (possibly as a result of oculomotor error) showed no reliable effect. The contextually inconsistent preview also gave rise to inflated fixation times when the target was fixated, but this effect could be attributed to a lack of orthographic overlap, rather than any extraction of parafoveal meaning.

While plausibility-related parafoveal-on-foveal effects remain controversial, it is widely accepted that the plausibility of words within a sentence can have an immediate impact on the duration of fixations falling on the word. For example, Rayner, Warren, Juhasz and Liversedge (2004) presented participants with a series of sentences in which a critical noun was either plausible (likely), implausible (unlikely) or anomalous (inappropriate), given the preceding sentence context. They found that the anomalous words had an immediate impact on gaze duration on the target word, while effects of implausibility were reflected only in later measures. Interestingly, they also reported a plausibility-related parafoveal-on-foveal effect, with gaze duration on the word preceding the anomalous one being 17ms and 14ms longer than in the control and implausible conditions, respectively. The authors, however, attribute this effect to oculomotor error.

Whatever one thinks about the interpretation of apparent semantic parafoveal-on-foveal effects, it is clear that manipulating the plausibility of a word can produce robust effects on the reading pattern when that word is fixated; indeed, there is additional evidence to suggest that effects of severe implausibility and anomaly can be reflected in measures as early as the very first fixation on a word (e.g., Staub, Rayner, Pollatsek, Hyönä & Majewski, 2007; Warren & McConnell, 2007). Given the clear and immediate nature of plausibility effects on target word processing, they provide an ideal manipulation for investigating whether or not parafoveal word meaning becomes activated prior to direct fixation.

This study therefore presented participants with sentences in which a critical word (n+1) was manipulated prior to fixation, such that it was initially previewed as a word that was either (a) identical, (b) different but plausible in terms of the preceding sentence context, (c) anomalous within the preceding sentence context, or (d) an illegal nonword. Once the eye passed an invisible boundary located before this critical word, all previews were replaced with the target word.

If meaning is extracted from the parafovea, it would be expected that an anomalous preview should exert an immediate impact on word n+1 fixations compared to the plausible preview condition. Conversely, if the meaning of the parafoveal word is not extracted while fixating word n, then plausible and anomalous previews should both produce the same cost, as a result of their lack of orthographic overlap with the target. An illegal nonword served as a baseline against which the magnitude of preview benefit could be judged.

5.2.1. Method

5.2.1.1. Participants

Twenty-eight native English speakers with normal or corrected to normal vision took part for course credits or £5 payment.

5.2.1.2. Materials and Design

Ninety-six experimental sentences were constructed. Each sentence contained a critical word pair comprising a 6-letter verb (word n) followed by a 6- or 7-letter noun (word n+1). To facilitate processing, word n was always high frequency (M=135, SD=90 occurrences per million, by Kuçera & Francis, 1967). Word n+1 was assigned one of four pre-fixation previews, all chosen to be very low in predictability. As can be seen from Figure 5.1 below, the previews were either: identical (e.g. “dinner” – the word changed to itself), plausible (e.g., “coffee” – an alternative that fitted the preceding context), anomalous (e.g. “caught” - a word that produced a semantic or grammatical violation), or an illegal nonword (e.g., "fumeio" – a letter string containing combinations not found in the English dictionary, in this case "eio"). The frequencies of these three preview words did not differ (all per million with standard deviation in parentheses: M=132 (176), M=144(203) and M=140(180), respectively; all *t*s<1). To allow each sentence to be displayed on a single line of the CRT display, sentence length ranged between 78 and 92 characters.

Previews were displayed until the eye passed an invisible boundary located prior to the space before word n+1, shown below with a "|". When the eye crossed this boundary, the identical preview was then displayed.

```
                    n              n+1
I) The mother was making| (dinner) dinner in the kitchen for her two
children and her husband.
P) The mother was making| (coffee) dinner in the kitchen for her two
children and her husband.
A) The mother was making| (caught) dinner in the kitchen for her two
children and her husband.
N) The mother was making| (fumeio) dinner in the kitchen for her two
children and her husband.
```

Figure 5.1. *Example item in each of the 4 parafoveal preview conditions: identical (I), plausible (P), anomalous (A) and nonword (N). Parafoveal previews are presented in parentheses, while the target word (n+1) is underlined. The boundary location is denoted by the symbol: "|".*

Twelve participants, who did not take part in the eye tracking experiment, provided plausibility ratings. Participants rated all three versions of each sentence up to and including word n+1 (illegal letter strings were not included). They used a rating scale from 1 (low) to 7 (high) plausibility. Additionally, they could use "U" instead of providing a numerical rating if they felt the sentences were ungrammatical; "U" scores were coded as 0 for purposes of analysis. The identical and plausible fragments were both rated as highly plausible (means=6.3 and 6.2, respectively), with no significant difference between these conditions ($t(95)=.85$, $p=.40$). The mean rating for the anomalous condition was 1.0 which differed significantly from both the identical ($t(95)=44.63$, $p<.001$), and plausible conditions ($t(95)=45.90$, $p<.001$).

Cloze task predictability was assessed with an additional 12 participants. Results from this confirmed that word n+1 was always of a very low predictability. The identical and plausible words were both correctly predicted on less than 3% of occasions, while the anomalous words were never correctly predicted.

Four counterbalanced item files were constructed. Each participant experienced all preview conditions across an equal number of items, but saw only one version of each item. The particular allocations of items to files and participants to files were treated as between-groups dummy variables in the following analyses.

To ensure normal reading for comprehension, 20% of the experimental items were followed by a comprehension question. In addition, a further 19 similar items were constructed as filler items. Eight separate practice items preceded the experimental items; half of these were accompanied by a comprehension question.

5.2.1.3. Apparatus

Identically with Experiment 2, participant's eye movements were recorded using the Dr Bouis eye-tracker. Full details of the apparatus can be found in Chapter 4, Section 4.2.1.3.

5.2.1.4. Procedure

The procedure was exactly the same as Experiment 2, detailed in Section 4.2.1.4.

5.2.2. Results and Discussion

For purposes of analysis, three zones were defined in each of the experimental items: one corresponding to each of the words n, n+1 (zones 1, 2, respectively), and a 3-word "spillover" region (zone 3). Fixations falling on the space preceding each of these were also considered to have fallen into the relevant region.

Zones: 1 2 3

```
The mother was| making| dinner| in the kitchen| for her two children and her
husband.
```

A number of fixation time measures are reported. These include: first, single and last fixation durations, gaze duration, go-past time and first-pass re-reading time. Saccadic measures include first landing positions and skipping probability. Please refer to Section 4.2.2 for a full description of each of these measures.

A one-way analysis of variance (ANOVA) was conducted on each of the above measures for Zones 1 to 3. Participants (F1) and items (F2) were treated as random variables and file was treated as a between groups dummy factor in both analyses.

Participants achieved an overall accuracy rate of 86%, suggesting they had read the sentence carefully.

5.2.2.1. Effects of Word N+1 Preview on Word N

As can be seen in Table 5.1, the probability of fixating word n did not vary across conditions (both Fs<1). The cumulative durational measures also failed to show any evidence that the nature of word n+1 modulated inspection times on word n (all Fs<1.3). Individual fixation duration measures did, however, show a pattern of shorter inspection times when word n+1 preview was anomalous rather than plausible (i.e., either an identical or plausible preview word[26]), although this only approached significance in last fixation duration (last fixation duration: $F1(3,72)=2.81$, $p<.05$; $F2(3,276)=2.05$, $p=.11$; first fixation duration: both $Fs<1.3$; single fixation duration: $F1<1$; $F2(3,276)=1.95$; $p=.12$, respectively). In this measure, pairwise comparisons showed no significant difference between the identical and plausible conditions ($F1(1,24)=2.07$, $p=.16$, $F2(1,92)=2.21$, $p=.14$), but when these two conditions were combined and compared to the anomalous condition, fixation durations were reliably shorter when an anomalous word was present in the parafovea ($F1(1,24)=5.30$, $p<.05$, $F2(1,92)=4.29$, $p<.05$), suggesting that the nature of this preview attracted attention from word n. The same trend emerged when an illegal nonword fell to the right of fixation, although the difference between the combined identical and plausible conditions and the illegal nonword condition was only reliable by subjects ($F1(1,24)=4.54$, $p<.05$, $F2(1,92)=1.10$, $p=.30$).

[26] Since the display change will not have occurred while fixating word n, both the identical and plausible parafoveal words can be considered equivalent at this stage.

Table 5.1. *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for Word N.*

	Identical	**Plausible**	**Anomalous**	**Illegal**
First Fix	257	263	255	258
Last Fix	257	264	253	253
Single Fix	264	273	261	266
Gaze	274	280	278	268
Go-Past	301	305	305	296
Re-Reading	27	24	28	28
Skip Prob	8	8	8	10
Landing	2.99	3.03	2.75	2.83

This immediate speed-up with anomalous previews appears similar to the finding by Starr and Inhoff (2004) of a tendency toward shorter gaze duration on the pre-target word when the target preview was contextually inconsistent. Interestingly Rayner et al (2004) reported an effect of parafoveal anomaly on word n in the opposite direction, with longer fixations when there was an anomalous word to the right of fixation. They concluded this must result from mislocated fixations, with the reader staying and processing word n+1 from a sub-optimal parafoveal location. But a reduction in fixation duration does not permit the same interpretation. These results and the pattern apparent in Starr and Inhoff's contextually inappropriate condition appear more consistent with Kennedy's (1998, 2000) attractor hypothesis, in which something unexpected in the periphery attracts attention, resulting in shorter fixation durations on the preceding word. Parafoveal-on-foveal effects of this sort have frequently been reported as a consequence of orthographic peculiarities to the right of fixation; however, here we see it, to an equivalent extent, as a consequence of meaning.

Word n+1 preview also appeared to modulate first landing position within word n ($F1(3,72)=3.55$, $p<.05$; $F2(3,276)=3.64$, $p<.05$). As can be seen in Table 5.1, there was no difference in first landing position on word n when the parafoveal word was either an identical or a plausible word (both $Fs<1$). However, the location of the first fixation when there was an anomalous word previewed to the right of fixation was significantly earlier compared to when the parafoveal word was either identical or a plausible preview (anomalous vs. identical and plausible combined: $F1(1,24)=7.80$, $p<.05$; $F2(1,92)=11.28$, $p<.01$). A similar trend of earlier first landing positions also emerged with an illegal nonword preview to the right of fixation compared with the identical or plausible conditions, although this effect was marginal (illegal nonword vs. identical and plausible combined: $F1(1,24)=3.82$, $p=.06$; $F2(1,92)=3.69$, $p=.06$). Decisions on where to fixate within word n must be made while fixating word n-1, so such an effect would appear to implicate a word n+2 preview effect.

Overall, this pattern of effects suggests that if the anomaly or illegality is detected on word n-1, a more cautious reading style is adopted. This contrasts with an apparent attraction mechanism engaged when the anomaly or irregularity is detected while fixating word n (see above). While the pattern of effects on word n indicates that anomalies and orthographic illegalities can be detected prior to fixation, the results presented here should be treated with a degree of caution since effect sizes were small, occasionally unreliable, and only reflected in a subset of measures. It is clear that replication is essential before strong conclusions can be drawn regarding the existence of plausibility-related parafoveal-on-foveal effects.

5.2.2.2. Word N+1 Preview Effects on Word N+1

The probability of skipping word n+1 was unaffected by prior preview ($F1(3,72)=1.50$, $p=.22$; $F2(3,276)=1.70$, $p=.16$), as was first landing position within it ($F1(3,72)=2.06$, $p=.11$; $F2(3,276)=1.85$, $p=.14$). There was, however, a consistent effect of prior preview on first fixation duration ($F1(3,72)=7.43$, $p<.001$; $F2(3,276)=5.81$, $p<.01$), single fixation duration ($F1(3,72)=5.86$, $p<.01$; $F2(3,276)=4.67$, $p<.01$), last fixation duration ($F1(3,72)=4.69$, $p<.01$; $F2(3,276)=3.5$, $p<.05$), gaze duration ($F1(3,72)=8.87$, $p<.001$; $F2(3,276)=8.71$, $p<.001$) and go-past time ($F1(3,72)=7.65$, $p<.001$; $F2(3,276)=9.29$, $p<.001$). But since the effect of preview was only significant by-items in first pass re-reading time ($F1(3,72)=1.82$, $p=.15$; $F2(3,276)=2.96$, $p<.05$) this appears to indicate that participants opted to refixate word n+1 rather than to consistently regress from it as a consequence of variation in preview. As can be seen from Table 5.2, the longest durations were associated with words that had previously received an illegal nonword preview, followed by anomalous then plausible previews, with identical previews associated with the shortest durations.

Table 5.2. *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for Word N+1.*

	Identical	Plausible	Anomalous	Illegal
First Fix	267	277	281	287
Last Fix	267	275	277	283
Single Fix	274	284	287	297
Gaze	298	299	313	331
Go-Past	331	345	370	389
Re-Reading	33	47	57	58
Skip Prob	5	6	7	4
Landing	3.21	3.29	3.13	3.04

Pairwise comparisons showed the expected orthographic preview effect, with all measures showing an increased inspection time following an illegal nonword preview compared to an identical preview (first fixation duration: $F1(1,24)=47.02$, $p<.001$; $F2(1,92)=18.06$, $p<.001$; single fixation duration: $F1(1,24)=38.75$, $p<.001$; $F2(1,92)=19.34$, $p<.001$; last fixation duration: $F1(1,24)=21.87$, $p<.001$; $F2(1,92)=14.97$, $p<.001$; gaze duration: $F1(1,24)=23.00$, $p<.001$; $F2(1,92)=23.68$, $p<.001$ and go-past time: $F1(1,24)=15.69$, $p<.001$; $F2(1,92)=26.67$, $p<.001$).

Readers also appear to have noticed the word change from a plausible preview to the target word, with a reliable increase in first fixation duration ($F1(1,24)=5.66$, $p<.05$; $F2(1,92)=5.00$, $p<.05$). Similar trends were also apparent in single fixation duration, last fixation duration and go-past time, although they failed to achieve statistical significance (single fixation duration: $F1(1,24)=3.16$, $p=.08$; $F2(1,92)=3.90$, $p=<.05$; last fixation duration: $F1(1,24)=3.22$, $p=.08$; $F2(1,92)=2.62$, $p=.11$ and go-past time: $F1(1,24)=3.70$, $p=.06$; $F2(1,92)=1.73$, $p=.19$); while there was no indication of an effect in gaze duration (both Fs <1). The trend in go-past time but not gaze duration highlights a tendency to regress from word n+1 rather than to refixate it following a plausible compared to an identical preview (first pass re-reading time: $F1(1,24)=3.19$, $p=.08$; $F2(1,92)=3.02$, $p=.08$; number of fixations: both Fs<1). Overall, this pattern suggests that the change from a different, though plausible, word was noticed immediately. It is not clear, however, whether this is an effect of meaning change, since it could equally be a consequence of the lack of orthographic overlap between the plausible preview and target.

A test of the effect of meaning can, however, be found in the contrast between plausible and anomalous previews, since both involve a change in orthography. The results here suggest that the meaning of word n+1 was indeed extracted while fixating word n. While first, single and last fixation durations showed no indication of an increased cost of anomaly (all $Fs<1$), the 14ms increase in gaze duration was significant by-subjects and approached significance by-items ($F1(1,24)=4.78$, $p<.05$; $F2(1,92)=2.99$, $p=.08$) and with regressions taken into account, the 25ms increase in go-past time was significant by-subjects and close to significant by-items ($F1(1,24)=5.88$, $p<.05$; $F2(1,92)=3.51$, $p=.06$).

As can be seen in Table 5.3, an increase in go-past time following anomalous previews also arose in the spillover region (see below). Combining word n+1 and the spillover regions, the difference in go-past between the plausible (898ms) and anomalous (946ms) conditions was significant by both subjects and items ($F1(1,24)=7.29$, $p<.05$; $F2(1,92)=6.17$, $p<.05$). This effect – an immediate and robust slowing following an anomalous preview – is the sort of effect found, for example, by Rayner et al (2004), who suggest that anomalous words “hit the reader over the head” (p. 1297). It is apparent, however, from these results that the genesis of this effect can also be parafoveal, with the reader detecting anomaly far earlier than previously thought.

These results strongly suggest that parafoveal preview effects are not limited to the extraction of orthographic and phonological features, but that higher-level linguistic processing can be engaged when previewing words to the right of fixation, and when the input changes, as happened here, prior processing of this sort can interfere with later comprehension.

Proponents of serial models, such as the E-Z Reader model, might suggest that these results stem from a quick succession of events all taking place while the eye is still directed at word n. It might be that word n receives full lexical access, allowing an attention shift to word n+1 which in turn also receives full lexical access and semantic interpretation; all prior to word n+1 being fixated. The plausibility of such an interpretation will be returned to in the General Discussion for this experiment.

5.2.2.3. Word N+1 Preview Effects in the Spillover Region:

An effect of prior preview was apparent in first landing position in the spillover region ($F1(3,72)=3.19$, $p<.05$; $F2(3,276)=2.58$, $p=.05$). As the pattern of means in Table 5.3 suggests, there was no difference between first landing positions when word n+1 had received identical, plausible or nonword previews (all $Fs<1$); however, earlier first landings were observed following an anomalous preview (identical vs. anomalous: $F1(1,24)=7.03$, $p<.05$; $F2(1,92)=7.40$, $p<.01$; plausible vs. anomalous: $F1(1,24)=3.79$, $p=.06$; $F2(1.92)=3.94$, $p<.05$ and illegal nonword vs. anomalous $F1(1,24)=4.01$, $p=.05$; $F2(1.92)=4.03$, $p<.05$). Prior anomalous previews appear to have prompted a more cautious reading strategy with subsequent fixations falling earlier within the spillover region.

Table 5.3. *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for the Spillover Region.*

	Identical	**Plausible**	**Anomalous**	**Illegal**
First Fix	256	254	254	255
Gaze	265	280	286	285
Go-Past	505	553	576	541
Re-Reading	41	73	90	55
Landing	4.41	4.32	4.01	4.31

First fixation duration and gaze duration were unaffected by word n+1 preview (both Fs<1 and $F1(3,72)=1.56$, $p=.21$; $F2(3,276)=1.62$, $p=.18$, respectively). There was, however, a highly significant effect of word n+1 preview in go-past time ($F1(3,72)=6.98$, $p<.001$; $F2(3,276)=10.48$, $p<.001$) and unsurprisingly, therefore also in first pass re-reading time ($F1(3,72)=4.40$, $p<.01$; $F2(3,276)=6.42$, $p<.001$).

Pairwise comparisons show that go-past time was significantly longer if word n+1 had received an illegal nonword preview rather than an identical one ($F1(1,24)=7.65$, $p<.05$; $F2(1,92)=5.99$, $p<.05$). But with the equivalent increase in first pass re-reading time nonsignificant ($F1(1,24)=1.25$, $p=.27$; $F2(1,92)=1.45$, $p=.23$) this effect appears to have been driven by a composite of extra fixations in the spillover region following a nonword preview (2.20 vs. 2.07 fixations; $F1(1,24)=6.51$, $p<.05$; $F2(1,92)=4.25$, $p<.05$) combined with a nonsignificant tendency to regress more.

As will be recalled from Chapter 4, such a finding is difficult to reconcile with serial models, such as E-Z Reader, which postulate that orthographic extraction occurs during the first stage of lexical processing on a word (L1), since it is the completion of this stage that triggers a saccadic program to be initiated to move the eyes to the next word. Spillover effects should never, therefore, be orthographic in nature, unless of course, they are the result of a mislocated fixation. For example, the L1 stage of lexical processing might not have been reached on word n+1, making it the target of an upcoming fixation. However, error in the oculomotor system might cause the eye to overshoot word n+1, resulting in an erroneous fixation on word n+2. Potentially, this scenario, if followed by a stay and process response to the error, might allow the expression of an effect of orthographic illegality to spill over into the duration of fixations falling on word n+2.

While the word change involved in the plausible condition produced increases in go-past and first pass re-reading times compared to the identical condition ($F1(1,24)=20.68$, $p<.001$; $F2(1,92)=17.53$, $p<.001$ and $F1(1,24)=5.54$,

$p<.05$; $F2(1,92)=8.30$, $p<.01$, respectively), there was no reliable effect when anomaly was compared to plausible in these measures ($F1(1,24)=1.70$, $p=.21$; $F2(1,92)=2.47$, $p=.12$ and $F1(1,24)=1.56$, $p=.22$; $F2(1,92)=1.32$, $p=.25$, respectively). Somewhat surprisingly, spillover effects related to a change of orthographic form from preview to direct fixation appear to show a longer time course than effects of parafoveal meaning. The latter exerts a more immediate impact, mostly reflected in cumulative durational measures on word n+1.

5.2.3. General Discussion of Experiment 4

This study set out to investigate whether preview benefit is restricted to the orthographic and phonological properties of a parafoveal word or whether meaning can also be extracted before a word is directly fixated. It was suggested that the inconsistent results in the semantic preview literature may be related to the use of semantically related previews, with the word change interfering with target processing, rather than facilitating the recognition process. By varying the plausibility of previews, two important semantic effects were revealed.

First, there was some suggestion that an anomaly in the parafovea can attract attention, resulting in shorter last fixation durations on pre-target words. While others have reported anomaly-related parafoveal-on-foveal effects (Kennedy, Murray & Boissiere, 2004; Murray, 1998; Murray & Rowan, 1998; Rayner et al, 2004; Starr & Inhoff, 2004), this appears to be the first study to find an effect that cannot be attributed to a mislocated fixation. If word n had been erroneously fixated instead of word n+1, triggering a stay and process response, then the pattern of effects on word n should mimic those seen on word n+1, with increased inspection time for anomalous previews, not reduced ones as seen here. As Liversedge, Paterson and Pickering (1998) point out, when faced with difficulty the reader has three options: (a) stay and resolve the problem, (b) make a regression, or (c) proceed, in anticipation that later words will help resolve the difficulty. It seems that here readers had a tendency to opt for the latter option, which as Liversedge et al point out, can result in reduced fixation durations.

Second, word n+1 preview effects were influenced by the plausibility of the preview. This effect cannot be explained in terms of orthographic overlap since both the plausible and anomalous previews differed from the target. The nature of the effect was distinctly different depending on whether a plausible or anomalous preview was employed, with anomalous previews exerting a more immediate and robust effect on word n+1 viewing times. Since these previews were only available prior to word n+1 receiving a direct fixation, the results clearly provide evidence for the extraction of meaning from a word in the parafovea.

Given the form of the word n+1 effects found here, it seems that the approach of looking for meaning effects using semantic associates is likely flawed. Even if there is some semantic facilitation from the associated word preview, the magnitude of this would appear to be more than outweighed by the interference generated by a word change, with these results clearly indicating a word change cost. This sort of combination of facilitation and inhibition might possibly explain why some studies fail to uncover semantic preview benefit using semantic associates (e.g., Rayner et al, 1986; 2014; Schotter, 2013) while others have been successful (e.g., Hohenstein & Kliegl, 2013; Schotter's synonymous condition; 2013), with the degree of correspondence between preview and target word meaning responsible for whether or not semantic preview benefits emerge. As outlined in the Introduction, Schotter (2013) has recently published a correlational analysis that appears to support such a proposal, with a negative relationship between target word inspection time and the extent of meaning correspondence with the preview.

The increased inspection time following a nonword preview compared to an unrelated real word also points, more generally, to the potential role of interference effects in gaze contingent change experiments which seek to uncover a 'preview benefit'. Hohenstein and Kliegl (2013) and White et al (2008) also obtained a similar pattern of effects, with longer inspection times following nonword compared to real word previews (c.f., Starr & Inhoff, 2004 who only observed a negligible, non-significant difference). Unlike the present study, Hohenstein and Kliegl (2013) and White et al (2008) both employed pronounceable, rather than illegal, nonword previews. This suggests that the increased inspection time here associated with illegal nonwords is unlikely to have been driven solely by the peculiarity of the preview encouraging parafoveal pre-preprocessing. The origin of these effects appears to be far more complex than that. As Hohenstein and Kliegl (2013) note, inflated inspection times following nonword compared to real-word parafoveal previews could reflect attempts to lexically encode a nonword preview, which might, in turn, produce inhibition that persists once the eye lands on the target. This proposal therefore allows, via inhibitory processes, for target word inspection time to be inflated following nonword previews compared to real-word previews. By extension, they further acknowledge the possibility that all previews might carry a cost in terms of inhibition, with semantic associate previews potentially causing less interference than the preview of an unassociated word.

If one accepts the argument that the pattern of plausibility-related preview effects reported here was driven by differing levels of interference, one might have expected Hyönä and Häikiö (2005) to also have obtained differences in inspection time based on whether the preview contained a neutral or a curse word. As will be recalled from Chapter 2, Hyönä and Häikiö observed no difference between these two conditions. However, in the example provided, Hyönä and Häikiö's neutral stimulus was also implausible (e.g., *"In my opinion, any animal's* _penny_*…").* Thus, the implausibility of the neutral previews might have produced enough interference to mask any effects caused specifically by the curse words.

Reconceptualising the results of semantic preview benefit experiments in terms of plausibility-related preview costs potentially provides an explanation for the inconsistent results of semantic 'preview benefit' studies. For example, it will be recalled from Chapter 2, that White et al manipulated the second part of a compound noun, such that prior to fixation the preview was either identical to the target (*vanillasauce*), semantically related (*vanillamustard*) or unrelated to the target (*vanillapriest*), or was a nonword. They report obtaining a semantic preview benefit on the second constituent of the two-part compound word. While this effect was only reflected in a late measure, regression path duration[27], there was evidence to suggest that a semantically related preview reduced inspection times compared with a semantically unrelated preview. Whether these results relate to overall degree of correspondence between the preview and target is not clear, although based on the example item provided, this explanation appears unlikely since although *sauce* and *mustard* share some sematic features, they cannot really be considered synonymous. However, the contrast used to test the existence

[27] A measure that includes all fixations on the compound after the second constituent was first entered and before it was exited to the right.

of a semantic preview benefit was between the related and unrelated previews*,* where we see a clear difference in the plausibility of the previews, with *vanillapriest* being anomalous, while *vanillamustard* seems merely implausible. So, it could be argued that even when a closer correspondence in meaning does not exist between semantic associates, the severe implausibility or anomaly of an unrelated preview might also drive a difference that could masquerade as a semantic preview effect.

Whether or not one accepts the proposal that interference-based effects are implicated in the inconsistent results in the semantic preview benefit literature, the present research does provide further evidence suggesting that word meaning can be extracted prior to fixation. This research both compliments and extends previous research (e.g., Hohenstein & Kliegl, 2013; Schotter, 2013) by not only demonstrating that parafoveal word meaning can be extracted prior to fixation, but that this information can also undergo further higher-level integrative processing. This finding is significant since it suggests that the coupling between fixation location and attention is much looser than can currently be accounted for within the E-Z Reader model (Reichle et al, 2009).

Hohenstein et al (2013) and Schotter (2013) have both proposed an alternative explanation for the variability in results in the semantic preview literature. They suggest that differences in the depth of orthography between languages might explain why some studies find preview effects using semantic associates (e.g. Hohenstein et al, 2013), while others do not (e.g., Rayner et al, 1986; 2014). Depth of orthography refers to the degree of correspondence between a word's orthographic form and its phonology. Languages with shallow orthographies benefit from a closer correspondence, while deep orthographies tend to show some divergence. Since German has a relatively shallow orthography, less time should be required to complete orthographic and phonological processing compared with a language such as English, which has a comparatively deep orthography. Consequently, there may be enough time for a semantic associate to facilitate target word processing in German, but due to time constraints, only synonyms with similar meanings are able to facilitate target word processing in English.

Schotter further suggests that the depth of orthography hypothesis is theoretically compatible with the E-Z Reader model since it suggests that, in English, only the preliminary stages of semantic processing have time to accrue on a word prior to fixation. However, in languages benefitting from a shallow orthography, such as German, a quicker succession of orthographic, phonological and semantic pre-processing might be engaged and this increased efficiency should, according to Schotter, permit time for a semantic associate preview to facilitate target word processing. This theory, is therefore consistent with the general premise that attention only briefly precedes the eye and would consequently be compatible with the E-Z Reader model.

The patterns of results obtained here does not, however, appear to support such an explanation. The present experiment was conducted in English, which has a deep orthography. Therefore, with no correspondence in meaning, word n+1 meaning should not have been extracted and plausibility related preview effects should not have arisen. The present set of results is therefore difficult to reconcile with the depth of orthography hypothesis.

Alternatively, a proponent of serial word processing might attempt to explain the present data with the suggestion that meaning effects arose when attention moved to word n+1 following foveal identification of word n, but before the eye movement was executed. However, this would necessitate that there is enough of a lag between the shift of attention and the eye movement not only to enable the L1 stage of processing of word n+1 to be completed, resulting in a potential skip, but that there was also enough time for L2 to be completed, allowing meaning extraction to occur. But word n+1 was skipped rarely and no more often when it had been anomalous. In any case, it seems rather unlikely that all this processing could somehow be shoehorned into the time between the termination of lexical processing of word n and the execution of the saccade out of it.

Despite word n+1 skipping rates being unaffected by preview condition, the possibility that parafoveal processing may have advanced to the semantic (L2) stage after the labile stage of saccadic programming to that word (M1) had passed, committing a saccade to word n+1 to execution, cannot be discounted. Indeed, precisely such cases should give rise to semantic preview benefit being expressed on word n+1.

To test this hypothesis, Schotter, Reichle and Rayner (2014) recently ran a simulation using the E-Z Reader model to help determine how frequently lexical processing advanced to the L2 stage prior to fixation. They found that there was an 8% chance of lexical processing advancing to L2 on word n+1 using Schotter's (2013) synonymous materials, although this dropped to just 2% using less "optimal" materials.

One problem with these simulations, however, is that it cannot be assumed that achieving the L2 stage of lexical processing on word n+1, or time spent within that stage, will necessarily equate to a word n+1 semantic preview effect. According to E-Z Reader, if this stage is reached before a saccade to word n+1 is committed to action, then word n+1 will be skipped. These cases will not, therefore, be able to contribute towards the expression of a sematic preview effect on word n+1. The authors attempted to control for this possibility by using low frequency, 8-letter words in the position of word n+1, making them less likely candidates for rapid parafoveal identification and therefore less likely to be skipped. However, according to Rayner and McConkie (1976), 8-letter words are still skipped on over 18% of occasions. Given the frequency of skipping of 8-letter words, combined with the relatively small values reported by Schotter et al, it cannot be assumed that Schotter et al's simulations necessarily demonstrate the E-Z Reader model can account for semantic preview benefit. But of course, it should be relatively easy to control for this confound by restricting the analyses to those cases where the L2 stage of lexical processing had been reached but the nonlabile stage had not. However, to date, this scenario has not been tested.

Without running simulations using the present materials, and differentiating cases where pre-processing advanced to the L2 stage before or after completion of the M1 stage of saccadic programming, it is difficult to assess whether E-Z Reader could indeed account for the current sets of effects. However, the low predictability (<.03) and the length (M=6.5-letters) of word n+1 would appear to conspire against such rapid parafoveal identification. Thus, while the 10th version of the E-Z Reader model (Reichle et al, 2009) includes a module for post lexical integrative processing, and could therefore potentially account for the plausibility effects reported here, the time constraints imposed by lexical processing would appear to prevent this higher level module from ever becoming engaged with parafoveal words.

A final but important finding obtained from the present study relates to the continuing effect of the masking of word n+1 with an orthographically dissimilar preview, as shown in longer fixation durations in the spillover region. It will be recalled that this has been a recurrent effect in the studies reported in this thesis, being shown in Experiments 2 and 3, as well having been reported by researchers in other labs (e.g. Radach et al, 2013). While serial models can account for shorter fixation durations in the spillover region (e.g., Angele & Rayner, 2011) - since longer fixation durations on word n+1 following an invalid preview would allow more word n+2 preview benefit to accrue – the finding of a continuing increase in fixation duration cannot be afforded the same interpretation.

The most advanced model of the parallel variety, SWIFT (Schad & Engbert, 2012), would also almost certainly fail to capture the present pattern of effects, not least because a mechanism for higher-level processing is currently absent from the model. However, as Hohenstein and Kliegl (2013) note, semantic preview effects are certainly within the "theoretical spirit" of the model. Within SWIFT, since 'easy' to process words reach their maximum level of activation faster than 'difficult' to process words, 'easy' words will enter the decreasing level of activation phase faster, and since this phase is associated with an increasingly dilated span of attention, they will permit more advanced parafoveal extraction compared with comparatively 'difficult' words. The use of high frequency words in the position of word n should therefore have provided optimal conditions for parafoveal processing to reach a relatively advanced stage within SWIFT. However, as stated above, SWIFT is not currently advanced enough to determine whether plausibility related parafoveal effects might be possible within its parallel architecture.

Overall, the results of this study are difficult to reconcile with models of eye movement control that allow only strictly serial-sequential lexical processing. There are, however, two ways in which these effects could be accounted for by the E-Z Reader model. First, as has already been discussed, words n and n+1 could have been processed very quickly. But while this scenario is rather unlikely given the time constraints imposed by the model, it remains a possibility. Second, these effects might have arisen following a mislocated fixation (e.g., Drieghe et al, 2008). According to this argument, word n might, on occasion, have been erroneously fixated due to an undershoot of the intended target – word n+1. If a stay and process strategy was then engaged, word n+1 could be processed while fixating word n. Assuming word n+1 was then subsequently fixated, such a scenario could give rise – within a serial framework – to the parafoveal plausibility effects seen here. While there is currently little evidence in favour of a stay and process response to a mislocated fixation (see Chapter 6; Kennedy, 2008; c.f., Drieghe et al, 2008), or for the possibility that fast parafoveal processing can account for the expression of plausibility-related preview benefits, testing these two possibilities provided the motivation for Experiment 5.

5.3. EXPERIMENT 5

Experiment 5 was designed to investigate whether the plausibility-related preview effects obtained in Experiment 4 could have been caused by a sequence of events that are, at least potentially, compatible with serial models of eye movement control during reading.

The existence or otherwise of semantic preview benefit and plausibility related parafoveal-on-foveal effects has historically been taken as evidence either for or against parallel lexical processing during reading. The recent resurgence of research on this topic has, however, altered this theoretical landscape considerably. Shortly after positive evidence for semantic preview benefit filtered into the literature (Hohenstein & Kliegl, 2013; Schotter, 2013), so did a series of simulations indicating that the E-Z Reader model is capable of accounting for them via a quick succession of lexical processing (Schotter et al, 2014). As discussed above, whether or not these simulations necessarily settle the issue remains an open question.

Another open question is whether parafoveal plausibility effects might be attributed to oculomotor error. An undershoot of word n+1 will result in a fixation falling on word n and since this word falls before the boundary, a stay and process response to this mislocation will result in the preview rather than the target being lexically processed. For word n+1 to have been targeted implies, according to E-Z Reader, that word n had been fully processed prior to this erroneous fixation. Thus this scenario permits advanced lexical processing of word n+1 while fixating a fully processed word n.

The oculomotor response to a mislocated fixation is investigated more fully in Chapter 6, but there is another way we can determine whether the E-Z Reader model is capable of accounting for the results of Experiment 4. That is, to see whether the effects are still present when the word undergoing the plausibility manipulation is located at word n+2. Moving the location another word downstream should remove the possibility that the effects can be accounted for by a mislocated fixation, since an error spanning two word units is highly unlikely. It should also remove the possibility that the effects can be accounted for by fast parafoveal processing of the fixated and parafoveal words while fixating word n. As Schotter et al (2014) demonstrate, while E-Z Reader can simulate pre-processing of word n+2 while fixating word n, the effect sizes were “modest” and never advanced to the L2 stage of lexical processing – where semantic processing is assumed to take place in the model. Thus, while it could be argued that E-Z Reader is capable of accounting for higher-level preview effects related to word n+1, the same cannot be said for effects stemming from word n+2.

It will be recalled from Chapter 2 that Radach et al (2013) manipulated word n+2 preview with words that were either nonwords, or real-words, the latter of which were either predictable or unpredictable. They obtained evidence that a predictable word n+2 changing to an unpredictable word resulted in increased inspection times when word n+2 was subsequently fixated. Radach et al interpreted this as a lexical word n+2 effect, with readers extracting lexical information from predictable previews. However, some authors suggest that predictability has a very early influence on the word recognition process. For example, to account for their finding of an unorthodox predictability-based parafoveal-on-foveal effect[28], Kliegl, Nuthmann and Engbert (2006) proposed a 'memory retrieval' mechanism that can become activated prior to fixation. They suggest that predictions about upcoming words are implicitly generated, and if the parafoveal word is highly predictable, it can be processed during a fixation on the foveal word. The eyes are only 'attracted' to the next word if it cannot be guessed due to its low predictability. Thus, while Radach et al created optimal conditions for word n+2 pre-processing, it could also be argued that their effects reflect relatively early stages of parafoveal processing.

Manipulating the plausibility of word n+2 previews should provide a stronger test, since to know whether or not a word is plausible, the meaning of that word will need to be known. Obtaining plausibility-related effects from word n+2 would necessitate that word n+2 had been processed up to (and beyond) the L2 stage of lexical processing prior to the eye passing word n – a possibility that is not currently accommodated within the E-Z Reader model.

This present experimental design also allowed the testing of another hypothesis. As will be recalled from Chapter 2, the proponents of E-Z Reader 7

[28] That is, longer foveal inspection times for highly predictable parafoveal words compared to unpredictable words.

(Reichle et al, 2003) introduced a “pre-attentive visual processing” stage that – because it operates in parallel – allows parafoveal orthographic irregularities to be detected prior to fixation. This stage has subsequently been used to account for the finding that word n+2 irregularities can affect word n+1 targeting decisions (see e.g., Chapter 4; also: Angele & Rayner, 2011; Pynte et al, 2004; Radach et al, 2013; also see Angele and Rayner, 2011, for a short discussion). On the basis of this argument, real-word n+2 previews, since they are not orthographically irregular, should not be detected via the low level attentional scan, and should not, therefore, influence word n+1 target decisions.

Radach et al (2013) did not obtain any evidence suggesting that alternative word n+2 previews (predictable or unpredictable) influenced word n+1 targeting strategies, only that nonword previews did. But, given the early and robust effects of the anomalous previews seen in Experiment 4, these previews might provide a stronger test of whether alternative words in the position of word n+2 can influence word n+1 targeting strategies.

Experiment 5 employed the same base materials as Experiment 4, but in this case, the verb and noun were separated by a 4-letter adjective. The preview conditions in the two experiments remained the same, with identical, plausible, anomalous and nonword previews employed. However, rather than appearing in the position of word n+1, these now occupied the position of word n+2. The decision to include a 4-letter rather than a 3-letter word n+1 was based on the finding that word n+2 preview effects can occur with words of this length (see Chapter 4), and because words of a greater length would be expected to increase processing demand (4-letter words are fixated 16% more often than 3-letter words: Rayner & McConkie, 1976) and therefore should reduce the potential, according to the E-Z Reader model, for a double attention shift. To contrast conditions under which a double attention shift might occur (with word n+1 lexically identified) with those where it would be impossible, word n+1 was either visible prior to passing the invisible boundary located immediately after word n, or it received an invalid preview.

If the same pattern of effects arises on word n+2 in the present experiment as in Experiment 4, then this would strongly suggest that these effects were not caused by a mislocated fixation. It would also be very unlikely that they could be accounted for by a double attention shift within the current architecture of the E-Z Reader model (Schotter et al, 2013). Furthermore, any evidence that the anomalous word previews influence word n+1 targeting decisions (or durational measures) will mean that these effects cannot be accounted for with the sort of low level attentional scan postulated for E-Z Reader. The presence of higher-level plausibility-related effects would strongly suggest that attention is distributed to multiple words simultaneously and that fixation location and the locus of attention are not as tightly coupled as might be suggested by serial models of eye movement control.

5.3.1. Method

5.3.1.1. Participants

Fifty-six native English speakers with normal or corrected to normal vision took part for course credit or £5 payment.

5.3.1.2. Materials and Design

The ninety-six experimental items from Experiment 4 were used in the present study, but with an adjective inserted between the verb and noun. This resulted in a critical region that contained the following 3-word triplet: a high frequency 6-letter verb (word n), followed by a 4-letter high frequency adjective (word n+1) and then a high frequency 6- or 7-letter noun (word n+2). To encourage attention to stretch to word n+2 while fixating word n, word n+1 was always high frequency (M=857, SD=713; occurrences per million, estimated by Kuçera & Francis, 1967; words n and n+2 frequency values can be found in Section 5.2.1.2 of this Chapter). The insertion of the adjective resulted in sentences that ranged in length between 83 and 97 characters, including spaces, allowing them to occupy a single line of a CRT display.

The noun received one of four previews prior to passing the invisible boundary. These were identical with the previews employed in the previous experiment, and comprised: identical, plausible, anomalous and illegal previews (see Section 5.2.1.2 of this Chapter for the word frequency values and statistics). It is important to remember that the insertion of word n+1 means that the noun and its associated previews are now properties of word n+2. In addition to manipulating word n+2 preview, word n+1 (the 4-letter adjective) received either an identical preview or a nonword preview. Both preview changes were triggered in tandem as the eye passed an invisible boundary located immediately after word n. As can be seen in Figure 5.2 below, this resulted in a 2 (word n+1 preview) x4 (word n+2 preview) design with a total of 8 conditions.

N+1 Identical: n n+1 n+2

I)The mother was making| (some dinner) some dinner in the kitchen for her two children and her husband.

P)The mother was making| (some coffee) some dinner in the kitchen for her two children and her husband.

A)The mother was making| (some caught) some dinner in the kitchen for her two children and her husband.

N)The mother was making| (some fumeio) some dinner in the kitchen for her two children and her husband.

N+1 Denied: n n+1 n+2

I)The mother was making| (zmuc dinner) some dinner in the kitchen for her two children and her husband.

P)The mother was making| (zmuc coffee) some dinner in the kitchen for her two children and her husband.

A)The mother was making| (zmuc caught) some dinner in the kitchen for her two children and her husband.

I)The mother was making| (zmuc fumeio) some dinner in the kitchen for her two children and her husband.

Figure 5.2. *Example item in each of the eight possible preview conditions: identical (I), plausible (P), anomalous (A) and nonword (N). Parafoveal previews are presented in parentheses, while the target words (n+1 and n+2) are underlined. The boundary location is denoted by the symbol: "|".*

Twelve participants who did not take part in the eye tracking experiment rated the plausibility of the sentence fragments. All three versions of each sentence were rated up to and including word n+2 (illegal letter strings were not included). A rating scale from 1 (low) to 7 (high) plausibility was used.

Additionally, participants could use “U” instead of providing a numerical rating if they felt the sentences were ungrammatical; “U” scores were coded as 0 for purposes of analysis. While both identical and plausible fragments were rated as highly plausible (means = 6.4 and 6.2, respectively), the mean difference between the two conditions was marginally significant ($t(95)=2.02$, $p=.05$). This difference appeared to have been driven by 4 items, the removal of these eliminated any difference between these conditions ($t(95)=1.04$, $p=.30$). In the results that follow, critical results involving this specific contrast were re-run with the reduced item set, but since significance levels did not vary between these analyses and those employing the full dataset, all reported results are based on the full set. Unsurprisingly, the mean rating for the anomalous condition was extremely low (mean = 1.0) and differed significantly from both the identical ($t(95)=49.01$, $p<.001$), and plausible conditions ($t(95)=49.10$, $p<.001$).

Cloze task predictability ratings completed by an additional 12 participants confirmed that word n+2 was always of very low predictability. The identical and plausible words were correctly predicted on less than 10% of occasions, while the anomalous words were never correctly predicted.

Eight counterbalanced item files were constructed. Each participant experienced all preview conditions across an equal number of items, but saw only one version of each item. The particular allocations of items to files and participants to files were treated as between-groups dummy variables in the following analyses.

To ensure normal reading for comprehension, 20% of the experimental items were followed by a comprehension question. In addition, a further 19 similar items were constructed as filer items. Eight separate practice items preceded the experimental trials; half of these were accompanied by a comprehension question.

5.3.1.3. Apparatus and Procedure

This was identical with Experiment 2, detailed in Sections 4.2.1.3 and 4.2.1.4.

5.3.2. Results and Discussion

For purposes of analysis, four zones were defined in each of the experimental items, one corresponding to each of the words n, n+1 and n+2 (zones 1, 2 & 3, respectively), and a 3 word “spillover” region (zone 4). Fixations falling on the space preceding each of these regions were also considered to have fallen into the relevant region.

```
Zones:             1      2     3          4
The mother was| making| some| dinner| in the kitchen| for her two children […]
```

The same fixation time and saccadic measures as reported in Experiment 2 were used here.

A 2 (word n+1 preview) x 4 (word n+2 preview) analysis of variance (ANOVA) was conducted for each measure in zones 1 to 4. Participants (F1) and items (F2) were treated as random variables and file was treated as a between-groups dummy factor in both analyses.

Participants achieved an overall accuracy rate of 83%, suggesting they had read the sentences carefully.

5.3.2.1. Word N

Effects of Word N+1 Preview As can be seen in Table 5.4, there was no evidence for an orthographic parafoveal-on-foveal effect in either first fixation duration ($F1(1,48)=2.14$, $p=.15$; $F2(1,88)=1.62$, $p=.20$), single fixation duration ($F1<1$; $F2(1,88)=1.69$, $p=.19$), or last fixation duration ($F1(1,48)=1.72$; $p=.19$; $F2(1,88)=1.58$, $p=.21$). And while there was some indication of longer inspection times when the word had received an invalid preview in the cumulative measures, this was clearly not typical for all subjects (gaze duration: 292ms vs. 286ms: $F1(1,48)=2.75$, $p=.10$; $F2(1,88)=5.74$, $p<.05$; and go-past time: 267ms vs. 260ms: $F1(1,48)=1.75$, $p=.19$; $F2(1,88)=4.31$, $p<.05$).

There was no evidence to suggest that any increase in the cumulative measures was due to an increased tendency to regress if the parafoveal word had received an invalid preview (first pass re-reading time: both Fs<1). While landing position in word n was unaffected by word n+1 preview, a tendency for reduced skipping did exist, such that an invalid preview of word n+1 reduced the probability of skipping word n (12% vs. 10%: $F1(1,48)=3.41$, $p=.07$; $F2(1,88)=7.08$, $p<.01$).

Table 5.4. *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for Word N.*

	N+1 Identical				N+1 Invalid			
	N+2 Identical	N+2 Plaus	N+2 Anom	N+2 Illegal	N+2 Identical	N+2 Plaus	N+2 Anom	N+2 Illegal
First Fix	255	249	251	252	253	251	246	245
Last Fix	251	246	251	248	250	247	246	242
Single Fix	259	256	258	254	256	256	257	248
Gaze	266	259	262	254	269	261	271	265
Go-Past	283	288	289	283	293	294	293	291
Re-Read	17	30	27	29	24	33	22	25
Skip Prob	12	13	9	12	11	10	10	8
Landing	2.87	2.86	3.01	2.77	2.86	2.84	2.83	2.84

As discussed in Chapter 4, variation in skipping probability as a function of parafoveal illegality is not new, with previous studies obtaining evidence that these previews can either engage a more cautious reading strategy (encouraging a fixation on the preceding word; Pynte et al, 2004), or attract a fixation (triggering a skip of the preceding word; Angele & Rayner, Exp 1, 2011; Radach et al, 2013). The results of the present study clearly fall into the former class. While it is suggested that these results no longer present a challenge to the E-Z model with the inclusion of a low level attentional scan, it is argued that the divergent strategies readers appear to engage might. Without simulations, it is unknown whether the E-Z Reader model would be capable of accounting for either of these strategies – let alone both.

Effects of Word N+2 Preview There was no evidence to suggest that word n+2 preview influenced inspection time on word n (first fixation duration: $F1(3,144)=1.33$, $p=.27$; $F2(3,264)=1.46$, $p=.22$; single fixation duration: $F1(3,144)=1.94$, $p=.12$; $F2(3,264)=1.70$, $p=.17$; last fixation duration: $F1(3,144)=1.01$, $p=.37$; $F2(3,264)=1.64$; $p=.18$; gaze duration: both $Fs<1.1$; go-past time: both $Fs<1$; first pass re-reading time: $F1(3,144)=2.01$, $p=.11$; $F2(3,264)=2.06$, $p=.10$), or targeting decisions (first landing position: both $Fs<1$; skipping probability: $F1<1$: $F2(3,264)=1.76$, $p=.15$).

There was also no suggestion that word n+1 preview modulated any word n+2 parafoveal-on-foveal effects, with a clear absence of interactions between the two preview manipulations (skipping probability: $F1(3,144)=2.07$, $p=.11$; $F2(3,264)=1.76$, $p=.15$; all other $Fs<1.1$).

5.3.2.2. Word N+1

Effects of Word N+1 Preview While first landing positions on word n+1 were unaffected by their preview condition (both $Fs<1$), there was a reduced probability of skipping word n+1 if it had previously received an invalid preview (35% vs. 28%: $F1(1,48)=33.07$, $p<.001$; $F2(1,88)=21.01$, $p<.001$). As can be seen from Table 5.5, this effect of preview also extended to the durational measures, with inflated inspection time following an invalid preview (first fixation duration: 267ms vs. 252ms: $F1(1,48)=27.34$, $p<.001$; $F2(1,88)=30.23$, $p<.001$; single fixation duration: 269ms vs. 253ms: $F1(1,48)=22.03$, $p<.001$; $F2(1,88)=33.55$, $p<.001$; last fixation duration: 266ms vs. 250ms: $F1(1,48)=26.90$, $p<.001$; $F2(1,88)=31.83$, $p<.001$; gaze duration: 206ms vs. 173ms: $F1(1,48)=56.54$, $p<.001$; $F2(1,88)=53.28$, $p<.001$; go-past time: 237 ms vs. 189ms: $F1(1,48)=66.43$, $p<.001$; $F2(1,88)=60.70$, $p<.001$; and first pass re-reading time: 31ms vs. 16ms: $F1(1,48)=32.45$, $p<.001$; $F2(1,88)=22.35$; $p<.001$). Again, therefore, these results provide clear evidence for an orthographic word n+1 preview benefit.

Table 5.5. *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for Word N+1.*

	N+1 Identical				N+1 Invalid			
	N+2 Identical	N+2 Plaus	N+2 Anom	N+2 Illegal	N+2 Identical	N+2 Plaus	N+2 Anom	N+2 Illegal
First Fix	255	250	250	252	271	266	266	264
Last Fix	250	251	250	250	270	267	269	259
Single Fix	255	252	251	253	274	267	268	269
Gaze	183	176	165	166	205	210	204	205
Go-Past	198	192	184	182	235	247	233	234
Re-Read	15	16	19	16	31	37	29	28
Skip Prob	32	34	38	36	28	28	29	28
Landing	2.20	2.09	2.13	2.06	2.09	2.02	2.30	2.19

Effects of Word N+2 Preview There were no early effects of word n+2 preview in any of the durational measures (first and single fixation duration: all Fs<1; last fixation duration: both Fs<1.2; gaze duration: $F1(3,144)=1.59$; $p=.19$; $F2(3,264)=1.59$, $p=.19$; go-past time: $F1(3,144)=1.36$, $p=.26$; $F2(3,264)=1.30$, $p=.27$ and first pass re-reading time: both Fs<1). There was also no evidence that word n+2 preview influenced word n+1 skipping rates ($F1(3,144)=1.26$, $p=.29$; $F2(3,264)=1.50$, $p=.21$). Despite obtaining no main effect of word n+2 preview in first landing position ($F1(3,144)=1.51$, $p=.21$; $F2(3,264)=1.23$, $p=.30$), this measure did reveal a trend towards an interaction between the effects of the two preview manipulations ($F1(3,144)=2.17$, $p=.09$; $F2(3,264)=2.44$, $p=.06$); the nature of this interaction is shown below in Figure 5.3.

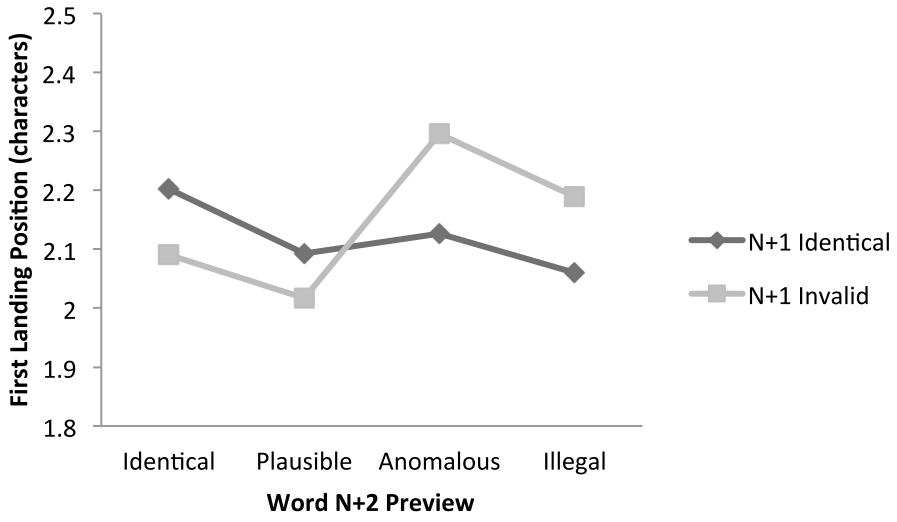


Figure 5.3. *Mean First Landing Position (character spaces) within Word N+1 as a Function of Word N+2 Preview for Each of the Two Word N+1 Preview Conditions*

As the Figure indicates, if word n+1 had been parafoveally available while fixating word n, there was no effect of word n+2 preview (both $Fs<1$). If, however, word n+1 had received an invalid preview prior to fixation, a significant effect of word n+2 preview did emerge ($F1(3,144)=2.98$, $p<.05$; $F2(3,264)=3.37$, $p<.05$). There was a tendency for first landing positions to fall closer to word n+2 if it had received an anomalous rather than an identical preview ($F1(1,48)=2.95$, $p=.09$; $F2(1,88)=7.67$, $p<.01$). Since a decision on where to fixate word n+1 must have been decided while fixating word n –

rendering the identical and plausible conditions qualitatively similar at this point – it is not surprising that while first landing positions did not differ between the identical and plausible conditions (both $Fs<1$), first fixations did land closer to word n+2 if it had received an anomalous rather than a plausible preview ($F1(1,48)=8.65$, $p<.01$; $F2(1,88)=6.77$, $p<.05$). These comparisons were nonsignificant when word n+2 had received an illegal nonword preview (identical vs. illegal: $F1(1,48)=1.35$, $p=.25$; $F2(1,88)=2.56$, $p=.11$; plausible vs. illegal: $F1(1,48)=4.35$, $p<.05$; $F2(1,88)=1.93$, $p=.16$).

These results appear to indicate that, provided word n+1 could not be pre-processed while fixating word n, attention proceeded to word n+2, where the anomaly was detected, resulting in an apparent attraction towards it. Given that an invalid preview of word n+1 should have, according to the E-Z Reader model, prevented attention from proceeding to word n+2, this model cannot easily account for these findings. However, since the qualifying interaction failed to achieve statistical significance, and this was combined with numerically small differences in first landing position, this appears to be a result that would benefit from replication.

There were no other interactions between the two preview manipulations in any of the measures (all $Fs<1$).

5.3.2.3. Word N+2

Effects of Word N+1 Preview Neither of the saccadic measures showed any delayed effect of word n+1 preview (first landing position and skipping probability: both $Fs<1$). There was also no delayed effect of word n+1 preview in first fixation duration (both $Fs<1.1$), single fixation duration (both $Fs<1$), last fixation duration ($F1(1,48)=2.75$, $p=.10$; $F2(1,88)=1.90$, $p=.17$), or gaze duration (both $Fs<1$). An invalid preview of word n+1 did, however, increase go-past time (334ms vs. 354ms: $F1(1,48)=9.33$, $p<.01$; $F2(1,88)=9.64$, $p<.01$) apparently driven by an increased tendency to regress following an invalid preview of word n+1 (42ms vs. 61ms: $F1(1,48)=18.44$, $p<.001$; $F2(1,88)=11.55$, $P<.01$).

Table 5.6. *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for Word N+2.*

	N+1 Identical				N+1 Invalid			
	N+2 Identical	N+2 Plaus	N+2 Anom	N+2 Illegal	N+2 Identical	N+2 Plaus	N+2 Anom	N+2 Illegal
First Fix	250	252	259	254	254	258	253	259
Last Fix	244	241	247	245	247	246	247	251
Single Fix	252	251	256	257	252	256	252	261
Gaze	282	290	304	292	292	299	291	291
Go-Past	314	334	350	337	364	356	351	344
Re-Read	32	44	46	45	72	57	61	53
Skip Prob	7	6	6	7	6	6	6	7
Landing	2.67	2.53	2.48	2.49	2.43	2.55	2.44	2.59

Since the effect was driven predominantly by regressions, proponents of serial models might suggest that these results were driven by an overshoot of word n+1 followed by a quick error correcting saccade back to that word. Given the recurring nature of this effect throughout this thesis, and the potential implication it has for serial models of eye movement control, this hypothesis was put to the test. To eliminate the possibility that this effect was driven by an accidental skip of word n+1, all cases in which word n+1 was skipped were removed from the analyses. The result clearly shows that the word n+1 preview spillover effect was still significant (*b*=-34.08, *SE*=12.85, *t*=-

2.65)[29]. This result suggests that the effect cannot be accounted for by an unintended skip followed by an immediate correction back to word n+1, since trials with skips were not included in the reduced dataset. Of course, the remaining possibility, that the effect was a product of a failed re-fixation of word n+1 stands, although given the length of word n+1 (4-letters) coupled with its high frequency, such a possibility seems extremely unlikely. Indeed, word n+1 received, on average, less than one fixation (0.83 fixations on average).

Effects of Word N+2 Preview Skipping probability and first landing position within word n+2 were unaffected by the preview manipulation (both Fs<1). While there were no main effects of word n+2 preview on first fixation duration (F1<1; F2(3,264)=1.27, p=.29), single fixation duration: (F1(3,144)=1.50, p=.21; F2(3,264)=1.74, p=.16), last fixation duration (both Fs<1), gaze duration (F1(3,144)=1.87, p=.14; F2(3,264)=1.13, p=.34), go-past time or first pass re-reading time (all Fs<1), there was a significant interaction

[29] Since this dataset involved missing data, analyses were run using a linear mixed effects model (LMM) including both subjects and items as random effects. Despite best efforts to attain a maximal random effects structure (Barr, Levy, Scheepers & Tily, 2013), the model failed to converge. Removing correlations and random intercepts did not aid convergence. As such the final model kept n+1 and n+2 preview as fixed effects (plus interaction), but – in accordance with Barr et al – the random effects structure was simplified in a principled way. The final model included: random intercepts and random slopes for n+1 preview, since this was the fixed effect of interest. The random slope for n+2 preview was omitted, so too was the random slope for its interaction with n+1 preview. The subject and item random effects structures were identical. An equivalent analysis was attempted with the slope for n+1 preview replaced with the slope for n+2, however, this model failed to converge. It should be noted that a by-subjects ANOVA on the reduced data set corroborates the reported result provided by the LMM (F1(1,48)=4.07; p<.05).

between the two preview manipulations in go-past time ($F1(3,144)=3.13$, $p<.05$; $F2(3,264)=2.92$, $p<.05$). This interaction can be seen in Figure 5.4. A numerically similar, but nonsignificant, interaction was also present in first pass re-reading time, suggesting that the effect in go-past was partially (but not entirely) driven by regressions ($F1(3,144)=2.08$, $p=.10$; $F2(3,264)=1.94$, $p=.12$).

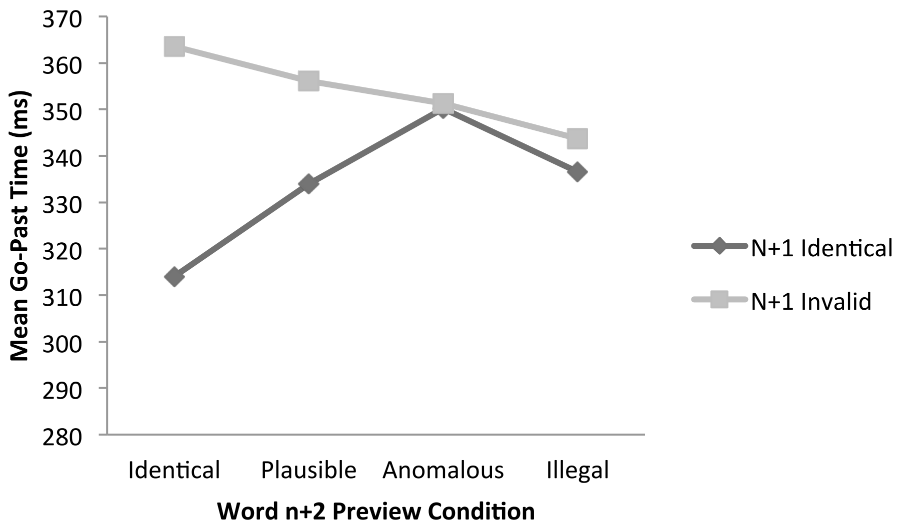


Figure 5.4. *Mean Go-Past Time (ms) on Word N+2 as a Function of Word N+2 Preview for Each of the Two Word N+1 Preview Conditions*

As can be seen from Figure 5.4, the word n+2 preview manipulation caused little variation in go-past time when word n+1 had received an invalid preview (both Fs<1). A significant effect of word n+2 preview was, however, observed following an identical preview of word n+1 ($F1(3,144)=3.91$, $p<.05$; $F2(3,264)=3.75$, $p<.05$).

All three alternative previews produced longer durations on the target word than when the preview had been identical, indicating the presence of an orthographic word n+2 preview benefit (identical vs. plausible: $F1(1,48)=3.92$, $p=.05$; $F2(1,88)=4.06$, $p<.05$; identical vs. anomalous: $F1(1,48)=10.89$, $p<.01$; $F2(1,88)=9.34$, $p<.01$ and identical vs. illegal nonword: $F1(1,48)=6.43$, $p<.05$; $F2(1,88)=6.59$, $p<.05$). This replicates earlier studies (e.g., Radach et al, 2007; 2013; Angele & Rayner, 2011) and further supports the existence of word n+2 preview benefit. Simulations involving the E-Z Reader model suggest that such effects no longer present a challenge to the serial perspective; but whether this is the case in the context of the present study will be returned to below and in the General Discussion of this chapter.

The pattern of the plausibility-related word n+2 preview effect was identical to that obtained on word n+1 in Experiment 4. Thus, the target words produced a similar pattern of means regardless of whether the target was located at word n+1 or word n+2. Identical previews were associated with the shortest durations, followed by the plausible previews, with the anomalous previews associated with the longest inspection time. The critical comparison with respect to the presence of a plausibility-driven word n+2 preview effect is the comparison between the plausible and anomalous conditions, which was not significant ($F1(1,48)=2.10$, $p=.15$; $F2(1,88)=1.32$, $p=.25$). Therefore, despite obtaining clear evidence for an orthographic word n+2 preview effect, there is only a numerical trend suggesting that word n+2 plausibility was extracted while fixating word n.

Interestingly, over this range, illegal nonword previews were no longer associated with the longest inspection times, with fixations following these previews showing a similar cost to that found with plausible previews (both $Fs<1$). Although the pattern of means suggests that anomalous previews interfered more with target word processing than nonword previews, this difference was not significant ($F1(1,48)=1.32$, $p=.25$; $F2<1$).

Therefore, the only statistical evidence for a preview effect on word n+2 was orthographic in nature. Only the pattern of means suggests the potential presence of a plausibility-related effect.

Supplementary analyses: A separate set of analyses were conducted to determine whether the orthographic word n+2 preview effect described above was restricted to cases where word n+1 was skipped or whether it extended to cases where word n+1 had been fixated. This is an important distinction since if these effects are only present following a skip of word n+1, they can easily be accounted for by the E-Z Reader model: Specifically, the completion of the L1 stage of lexical processing on word n triggers a saccade to the next word in text to be programmed. This saccade can, however, be cancelled and a new one to word n+2 programmed providing both the second stage of lexical processing on word n, and the first stage of lexical processing on word n+1 have both taken place prior to that saccade entering its nonlabile stage. This would be feasible given that the effect was only present when word n+1 was available for pre-processing. If time permits this sequence of events to occur, then attention can proceed once more onto word n+2, provided the second stage of lexical processing occurs on word n+1 first. Given that a time penalty will accompany the re-programming of a saccade from word n+1 to word n+2, it is possible that enough time will be available to allow word n+2 pre-processing prior to the eye exiting word n. Thus, word skipping in the E-Z Reader model could be associated with a double attention shift – a mechanism within the model responsible for accounting for orthographic word n+2 preview benefits (Schotter et al, 2014).

However, as can be seen from Figure 5.5, when first pass reading time was restricted to cases where word n+1 was (a) identical, and (b) fixated, the orthographic word n+2 preview benefit remained significant (all invalid n+2 previews combined vs. n+2 identical: SE=-21.47; b=9.67; t=-2.22)[30]. This finding replicates Radach et al (2013), who also reported an orthographic word n+2 preview effect on word n+2 when word n+1 was fixated.

[30] Again, since this dataset involved missing data, this analysis was conducted using an LMM including both subjects and items as random effects. Word n+2 preview (valid vs. invalid) was entered as a fixed effect. The random effects structure included random intercepts and slopes for word n+2 preview in both subject and item random effects structures.

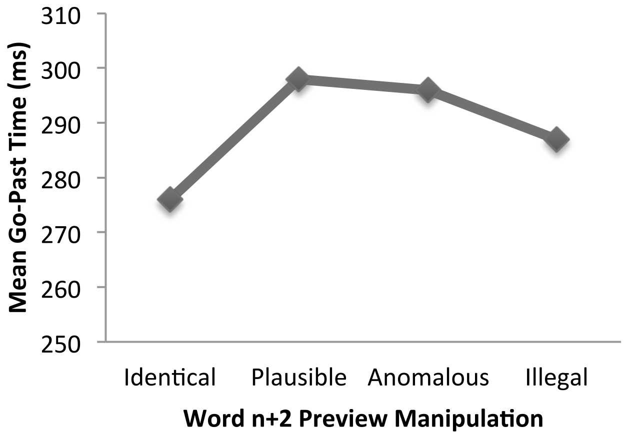


Figure 5.5. *Mean Go-Past Time (ms) on Word N+2 when Word N+1 was Identical and had been Fixated in Each of the 4 Word N+2 Preview Conditions*

It should be noted that this supplementary analysis does not preclude the possibility that the effect was driven by a double-attention shift. Specifically, word n+1 pre-processing might have been too slow for the saccade into it to be cancelled (i.e., L1 on word n+1 might have been reached during the nonlabile stage of saccadic programming), but it still might have been fast enough for lexical processing to complete on word n+1 and for attention to move forward onto word n+2, all prior to the eye passing word n. However, as Radach et al note, such cases should be exceedingly rare.

There were no other interactions between the word n+1 and n+2 preview manipulations in any measures (all *p*s>.10)

5.3.2.4. Spillover Region

Effects of Word N+1 Preview As can be seen from Table 5.7, any word n+1 spillover effects appear to have been confined to word n+2, with no evidence that they continued into the spillover region in any of the reported measures (all Fs<1).

Table 5.7. *Fixation Time Measures (ms) and First Landing Positions (character spaces) for the Spillover Region.*

	N+1 Identical				N+1 Invalid			
	N+2 Identical	**N+2 Plaus**	**N+2 Anom**	**N+2 Illegal**	**N+2 Identical**	**N+2 Plaus**	**N+2 Anom**	**N+2 Illegal**
First Fix	245	243	246	245	240	245	245	247
Gaze	456	448	451	456	462	444	452	454
Go-Past	498	496	493	494	509	495	494	496
Re-Read	42	48	42	38	47	51	42	42
Landing	4.54	4.31	4.54	4.37	4.42	4.49	4.20	4.53

Effects of Word N+2 Preview There were no delayed effects of word n+2 preview in any measure (gaze duration: $F1(3,144)=1.74$, $p=.16$; $F2(3,264)=1.23$, $p=.30$; all other Fs<1). While there was evidence for an interaction between the two preview manipulations in first landing position ($F1(3,144)=3.23$, $p<.05$; $F2(3,264)=2.67$, $p<.05$), as can be seen from Table 5.7, the differences were numerically small and follow up analyses failed to uncover a main effect of word n+2 preview regardless of whether word n+1 received an identical or invalid preview while fixating word n ($F1(3,144)=1.42$, p=.24; F2(3,264)=1.84, p=.14 and F1(3,144)=2.82, p=.08; F2(3,264)=2.07, p=.10, respectively).

There were no other interactions between the two preview manipulations in any other measure (all Fs<1).

5.3.3. General Discussion of Experiment 5

The primary aim of Experiment 5 was to further test the presence of plausibility-related preview effects in a paradigm where they cannot be attributed to either (a) a fast succession of lexical processing, or (b) a mislocated fixation. To this end, a four-letter adjective was inserted between word n and the target word undergoing the plausibility manipulation.

The plausibility manipulation employed here produced a similar pattern of effects on word n+2 to that seen in Experiment 4 when the plausibility manipulation was on word n+1. This pattern was, however, only present when word n+1 had been parafoveally available. As before, there was evidence to suggest that real-word previews had been parafoveally identified prior to fixation and that, upon fixation, the change in input from preview to target resulted in differing levels of interference, with anomalous previews being associated with a higher level of disruption than the plausible previews.

Despite the similarity in this pattern across the two experiments, the difference between the plausible and anomalous previews failed to achieve statistical significance. As such, it cannot be argued that plausibility-related preview benefits exist under conditions that are incompatible with serial models of eye movement control during reading. Indeed, one might argue that since the pattern was only apparent when word n+1 had been available for pre-processing, this supports the suggestion that these trends could have been caused by a quick, iterative succession of lexical processing.

While such an explanation is theoretically possible, it is unlikely that such a loose coupling between fixation location and attention could be accommodated within the current architecture of the E-Z Reader model. As Schotter et al (2014) have shown, E-Z Reader simulations are capable of accounting for orthographic but not semantic word n+2 preview effects. In fact, their simulations demonstrated that word n+2 pre-processing never advanced into the L2 stage wherein semantic information is extracted – a necessary pre-requisite for the word's plausibility to exert an effect. Given the numerical replication of the pattern of plausibility-related effects from Experiment 4 to Experiment 5, together with the importance of such a finding for models of eye movement control during reading, the pattern of means reported here clearly suggests the need for a follow-up study. This will be returned to below.

If we consider the pattern of preview effects on the critical word in both experiments, it is apparent that while there was a clear cost specifically associated with a nonword preview in Experiment 4, in this study the cost was very similar for all previews that were not identical with the target. This difference between the experiments could be related to the degree of confidence the reader had in ’perceiving’ the peculiarity in the periphery. If an illegal nonword occurs as word n+1, acuity should be reasonably high, resulting in a high level of confidence that the nonword does not conform to the usual orthographic rules. This may lead to the parafoveal nonword attracting attention, resulting in a high level of interference once the word is eventually fixated. In contrast, where the nonword is located two words downstream, although the peculiarity may have been parafoveally detected, confidence in the precise nature of that peculiarity may have been low due to acuity constraints, allowing that information to be more freely disregarded following direct fixation and therefore causing less interference.

Despite inconclusive evidence for a word n+2 plausibility-related preview effect, there was clear evidence that a change in the orthography of word n+2 from preview to direct fixation resulted in longer inspection times. Since this effect was restricted to cases where word n+1 had received a valid preview, it is possible that this effect was driven by a double attention shift. For the reasons outlined in the Results and Discussion, according to E-Z Reader, double attention shifts should be most prevalent following the skipping of word n+1. However, if word n+1 lexical processing is not advanced enough to cancel a saccade to that word by the time it was committed to action, this implies that word n+1 lexical processing was not progressing at speed. Therefore these cases are less likely to be associated with a secondary attention shift onto word n+2. Furthermore, upon fixation of word n+1, word n+2 will be visible, which means that any effect of the previously incorrect word n+2 preview is likely to be diminished in response to the now-correct parafoveal input from that word.

Within this architecture, therefore, word n+2 pre-processing should be more likely to occur or become further advanced following the skipping of word n+1, than when word n+1 is fixated. However, like Radach et al (2013), the present results suggest that word n+2 preview effects were still present following a fixation of word n+1. While this finding does not preclude the possibility that a double attention shift could account for orthographic word n+2 preview effects, it does reduce the probability, since it links them to cases where time constraints should act to prevent their routine occurrence.

It will be recalled that orthographic word n+2 preview effects were also reported in Experiment 3 of this thesis – a study that also employed 4-letter words in position n+1; the pattern of effects were, however, very different to those reported here. In Experiment 3, when word n+1 had received a valid preview, an invalid preview of word n+2 reduced, not increased, word n+2 inspection times. This difference could be associated with differences in overall reading strategies. In Experiment 3, an invalid preview of word n+2 had a tendency to increase inspection times on word n+1, which may have resulted in a trade-off in which the increased duration on word n+1 might have allowed more time to parafoveally process (a now visible) word n+2. Consequently, less time might have been required to process word n+2 once it was eventually fixated. In the present study, word n+2 preview did not modulate word n+1 inspection times, and therefore the same potential trade-off did not exist. This explanation is clearly speculative, and further research is required to help understand the complex nature of the effects of preview on reading strategies.

A secondary motivation for running Experiment 5 was to establish whether delayed parafoveal-on-foveal effects – that is, an effect of word n+2 preview expressed on word n+1 – exist for real-word as well as nonword n+2 previews. As will be recalled from the Introduction to this chapter, previous evidence for such modulation (e.g., Chapter 4; Kliegl et al, 2007; Radach et al 2013; Pynte et al, 2004) has been attributed by some to a low level attentional scan detecting upcoming peculiarity, causing the reader to adjust their reading strategies accordingly (e.g., Angele & Rayner, 2011). This study therefore provided the ideal opportunity to further investigate this possibility: if delayed parafoveal-on-foveal effects remain with real-word previews that are devoid of any orthographic peculiarities, then it is very unlikely that a low level attentional scan could account for such a pattern of effects.

Effects of word n+2 preview are often absent in durational measures but present in targeting measures (e.g., Angele & Rayner, Exp1, 2011; Pynte et al, 2004), which was also the case in the present experiment. Here, however, first landing position within word n+1 was closer to word n+2 following an anomalous preview compared to either an identical or plausible preview. There was also a trend for first landing position within word n+1 to fall closer to word n+2 when there had been a nonword preview than when it was either identical or plausible. Importantly, this was only the case when word n+1 had received an invalid preview. So within the context of the E-Z Reader model, these effects cannot be accounted for by a double attention shift. Nor can they be accounted for by a failed skip of word n+1 coupled with a stay and process response (e.g. Drieghe et al, 2008; see also Schotter et al, 2012), since the condition for word skipping within the model is that word n+1 is parafoveally identified – a condition that cannot be satisfied with a nonword preview of word n+1.

Also, since these effects were associated with an anomalous rather than orthographically illegal preview of word n+2, this means that they cannot have been driven by a low level attentional scan, since there was no orthographic irregularity. Angele and Rayner (2011) obtained a similar result, finding that nonword previews of both words n+1 and n+2 appeared to encouraged first landing position on word n+1 to fall closer to word n+2. In accounting for these effects, they suggested that the illegality of word n+2 might have attracted attention, a process they suggest could have been driven by a low level attentional scan. However, the same interpretation does not seem viable here, when word n+2 preview was a real-word. But the effect should be treated cautiously, since the qualifying interactions failed to achieve statistical significance and replication would certainly be desirable.

The word n+1 preview manipulation showed evidence for both localised and distributed effects. A standard word n+1 preview benefit was obtained with increased inspection time on word n+1 following a previously invalid preview. There was also some evidence to suggest that word n+1 illegality triggered a more cautious reading strategy, with an increased tendency to fixate word n when word n+1 was parafoveally unavailable. This effect is qualitatively similar to a result found in Experiment 2 and an effect reported by Pynte et al (2004) in which an illegal preview of word n+2 increased word n+1 fixation probability. It does, however, contrast with several other studies that have shown that remote orthographic peculiarity can attract fixations away from intervening words (e.g., Experiment 3 of this thesis; Angele & Rayner, Exp 1, 2011; Radach et al, 2013).

As previously discussed, proponents of the E-Z Reader model might suggest that such effects can be accounted for by the inclusion of the low level attentional scan in version 7 of the model (Reichle et al, 2003), but it is not clear how this could give rise to such contrasting effects. In any case, the recurrent nature of such effects would appear to call for the full implementation of this mechanism within the model to determine whether it is indeed capable of replicating such divergent findings.

The results here again showed that after passing word n+1, there was a spillover effect related to word n+1 preview, with longer durations following previously invalid previews. Orthographic spillover effects such as this have become a recurrent theme throughout this thesis and, as previously discussed, they appear to pose a significant challenge to the E-Z Reader model. The only way the model appears capable of accounting for orthographic spillover effects such as these is to assume that word n+1 was processed while the eye was positioned on word n+2, which could happen following an overshoot of word n+1 followed by a stay and process response. However, this scenario seems rather unlikely given that the effect remained significant following a fixation of word n+1. Thus these effects cannot be accounted for by an accidental skip of word n+1, which leaves one other possibility: that word n+2 was erroneously fixated following a failed refixation of word n+1. This latter explanation seems somewhat implausible, however, since there was little evidence to suggest that word n+1 required a second fixation.

To summarise, this study obtained three results that appear to present a challenge to serial models of eye movement control. First, there was evidence for a word n+2 preview benefit from orthographic overlap following a fixation on word n+1. Second, there was some suggestion that word n+1 preview modulated word n skipping rates. Finally, a word n+1 orthographic spillover effect was present even when word n+1 had been fixated.

The main purpose of this experiment was, however, to further investigate whether plausibility-related preview effects exist in an environment where their occurrence cannot be attributed to either (a) a quick succession of lexical processing, or (b) a mislocated fixation. Despite obtaining a strikingly similar pattern of plausibility-related preview effects on the target words of Experiments 4 and 5, the results of the present study were not statistically reliable in this respect. In order to determine whether the plausibility-related results of the present experiment failed to achieve significance due to either a lack of power or less-than optimal materials, a third variation of this study was undertaken.

5.4. EXPERIMENT 6

Experiment 6 was designed to create conditions that would be optimal for word n+2 pre-processing to occur. First, the length of word n+1 was shortened, becoming a 3-letter high frequency word. This would appear to increase the likelihood of finding an effect, since previous studies have typically failed to uncover positive evidence for word n+2 effects when word n+1 exceeds 3-letters (e.g., Angele et al, 2008; Rayner et al, 2007).

Secondly, since the plausibility-related preview effects obtained in Experiment 5 were restricted to cases where word n+1 had received a valid preview, to help increase power, word n+1 was always visible in the present experiment. While this could leave open the possibility that word n+2 preview effects might result from a double attention shift (see Chapter 2 and Angele et al, 2008), any evidence for plausibility-related word n+2 preview effects would still present a challenge to the E-Z Reader model, since, in its current instantiation, the model is incapable of predicting n+2 preview effects that extend beyond orthographic processing (Schotter et al, 2014).

5.4.1. Method

5.4.1.1. Participants

Forty native English speakers with normal or corrected to normal vision took part for course credits or £5 payment.

5.4.1.2. Materials and Design

The ninety-six experimental items from Experiment 4 were again used in the present study, but this time with either the determiner “the” (Kuçera & Francis frequency: M=69971/mil) or a three-letter high frequency word (M=1673/mil; SD=1906/mil) inserted between the verb and noun[31]. The critical region thus comprised: a high frequency 6-letter verb (word n), followed by a high

[31] To allow word n+1 to fit in between words n and n+2 without compromising the fluency of the sentence, the endings of 7 items (following word n+2) were re-written.

frequency 3-letter word (word n+1) then a high frequency 6- or 7-letter noun (word n+2). The change in word n+1 resulted in sentences that ranged from 82 to 96 character spaces. These were accommodated on a single line of the CRT display.

In the same manner as the preceding two experiments, the noun (word n+2) initially received one of four previews: identical, plausible, anomalous or illegal. Previews were identical with those used in the previous two experiments except that eight of the anomalous previews were substituted to ensure that word n+2 remained anomalous. Frequencies of these substituted words were held constant between this and the preceding experiments. All four previews changed to the target form of word n+2 once the eye passed an invisible boundary located immediately after word n. This resulted in a 2 (word n+1 type) x 4 (word n+2 preview) design with a total of 8 conditions, examples of which are provided in Figure 5.6 below:

N+1 determiner: n   n+1   n+2

I)The mother was making| the (dinner)dinner in the kitchen for her two children and her husband.

P)The mother was making| the (coffee)dinner in the kitchen for her two children and her husband.

A)The mother was making| the (caught)dinner in the kitchen for her two children and her husband.

N)The mother was making| the (fumeio)dinner in the kitchen for her two children and her husband.

N+1 high frequency word: n   n+1   n+2

I)The talented photographer showed| her (images)images to the paying client, who loved them.

P)The talented photographer showed| her (guests)images to the paying client, who loved them.

A)The talented photographer showed| her (minute)images to the paying client, who loved them.

N)The talented photographer showed| her (uiopnm)images to the paying client, who loved them.

Figure 5.6. *Example item in each of the 4 parafoveal preview conditions: identical (I), plausible (P), anomalous (A) and nonword (N). Parafoveal previews are presented in parentheses, while the target words (n+1 and n+2) are underlined. The boundary location is denoted by the symbol: "|".*

Plausibility ratings for these revised items were obtained from 12 participants who did not otherwise take part in this or the previous studies, following the same procedure used in Experiments 4 and 5. The anomalous words had a mean rating of just 0.9 and were rated as significantly less plausible than either the identical (M = 5.87; $t(95)=43.13$, $p<.001$) or plausible words (M = 5.55; $t(95)=45.02$, $p<.001$). Identical words were also rated as being significantly more plausible than plausible words ($t(95)=2.50$, $p<.05$). Since this difference was driven by a number of items, removing them from the analysis would have compromised statistical power. Consequently, in order to ensure that any difference obtained between these conditions was not driven by differences in plausibility, wherever an effect was observed in an ANOVA, the data were re-analysed using linear effects models (LMMs) entering plausibility as an additional continuous predictor variable.

Cloze task predictability ratings completed by an additional 12 participants confirmed that word n+2 was always of very low predictability. The identical and plausible words were correctly predicted on less than 6% of occasions, while the anomalous words were never correctly predicted.

Four counterbalanced item files were constructed. Each participant experienced all preview conditions across an equal number of items, but saw only one version of each item. The particular allocations of items to files and participants to files were treated as between-groups dummy variables in the following analyses.

To ensure normal reading for comprehension, 20% of the experimental items were followed by a comprehension question. In addition, a further 19 similar items were constructed as filler items. Eight separate practice items preceded the experimental items; half of these were accompanied by a comprehension question.

5.4.1.3. Apparatus and Procedure

These were identical with Experiments 2 to 5.

5.4.2. Results and Discussion

For purposes of analysis, four zones were defined in each of the experimental items: one corresponding to each of the words n, n+1 and n+2 (zones 1, 2 and 3, respectively), and a 3-word "spillover" region (zone 4). Fixations falling on the space preceding each of these were also considered to have fallen into the relevant region.

```
 Zones:            1      2       3         4
The mother was| making| the| dinner| in the kitchen| for her two children {…}
```

The same set of saccadic and fixation time measures used in the earlier experiments in this chapter are also reported here.

A 2 (word n+1 type) x 4 (word n+2 preview) analysis of variance (ANOVA) was conducted for each of the above measures for zones 1 to 4. Participants (F1) and items (F2) were treated as random variables and file was treated as a between groups dummy factor in both analyses.

Participants achieved an overall accuracy rate of 88%, suggesting they had read the sentences carefully.

It was clear that word n+1 type did not interact with word n+2 preview in any measure across any of the four regions: ***word n*** (last fixation duration: $F1(3,108)=2.56$, $p=.06$; $F2(3,264)=1.60$, $p=.19$; all other $ps>.14$); ***word n+1*** (skipping probability: $F1(3,108)=2.03$, $p=.11$; $F2(3,264)=2.35$, $p=.07$; all other $ps>.14$); ***word n+2*** (all $ps >.11$); ***spillover region*** (all $Fs <1$). For transparency, the effects of word n+1 type will be briefly summarised below, and then a more thorough discussion of the effects of the word n+2 preview manipulation will follow. For clarity, the means reported in Table 5.8 are listed again below in Tables 5.9 to 5.12 collapsed over the two types of word n+1.

5.4.2.1. Effects of Word N+1 Type

5.4.2.1.1. Word N: A trend towards a lexical parafoveal-on-foveal effect was obtained on word n: last fixation duration was 5ms shorter when the word to the right was an alternative high frequency word compared to a determiner; this was not however, typical of all items (243ms vs. 248ms: $F1(1,36)=4.81$, $p<.05$; $F2(1,88)=2.93$, $p=.09$). While it is acknowledged that this may only be driven by a subset of items, it should be noted that a similar effect was found in first pass re-reading time in Experiment 2. Taken together, these results suggest that the class of an upcoming word can influence reading strategies prior to fixation, and unlike orthodox parafoveal-on-foveal effects, these results cannot be explained by a mislocated fixation, since such an account would predict the opposite pattern of effects, with shorter times on the determiner.

As can be seen from Table 5.8, a similar trend was observed in first and single fixation duration, although this was not statistically reliable (first fixation duration: $F1(1,36)=2.09$, $P=.15$; $F2<1$; single fixation duration: $F1(1,36)=2.95$, $p=.09$; $F2(1,88)=1.73$, $p=.19$). Word n+1 type did not affect any other measure (skipping probability: $F1(1,36)=1.89$, $p=.17$; $Fs<1$; all other $Fs<1.1$).

Table 5.8. *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for Words N, N+1, N+2 and the Spillover Region.*

	N+1 Determiner				N+1 HF Word			
	N+2 Identical	N+2 Plausible	N+2 Anom	N+2 Illegal	N+2 Identical	N+2 Plausible	N+2 Anom	N+2 Illegal
Word N								
First Fix	249	254	252	248	246	249	245	249
Last Fix	245	253	249	244	246	246	234	247
Single Fix	253	259	256	250	248	254	245	254
Gaze	249	257	258	244	252	254	255	258
Go-Past	279	291	281	273	288	279	288	290
Re-Read	30	34	23	29	36	25	33	32
Skip Prob	12	11	11	13	10	12	10	11
Landing	2.66	2.91	2.84	2.90	2.86	2.94	2.71	2.79
Word N+1								
First Fix	230	226	229	230	251	235	241	250
Last Fix	232	226	228	229	249	235	242	247
Single Fix	230	227	229	231	251	237	243	251
Gaze	113	82	94	101	132	123	124	121
Go-Past	123	98	106	113	151	131	143	136
Re-Read	10	16	11	12	19	9	19	15
Skip Prob	53	64	61	57	46	49	50	53
Landing	1.31	1.35	1.31	1.19	1.47	1.35	1.52	1.49
Word N+2								
First Fix	259	260	271	259	261	264	265	258
Last Fix	249	249	257	248	250	247	252	249
Single Fix	261	264	276	261	262	266	265	256
Gaze	280	306	292	294	291	311	304	300
Go-Past	333	359	342	356	351	351	368	357
Re-Read	53	53	51	62	60	40	65	57
Skip Prob	7	5	7	5	6	5	6	6
Landing	2.91	2.78	2.88	2.99	2.51	2.54	2.41	2.38
Spillover								
First Fix	250	250	248	248	249	256	247	247
Gaze	491	470	483	470	470	469	470	470
Go-Past	556	548	556	536	526	521	530	508
Re-Read	66	79	72	66	56	53	60	38
Landing	4.55	4.58	4.62	4.78	4.40	4.48	4.44	4.52

5.4.2.1.2. Word N+1: With the exception of first pass re-reading time, there was a clear effect of word n+1 type on word n+1, with longer inspection times for alternative high frequency words compared to determiners in all measures (first fixation duration: 244ms vs. 229ms: $F1(1,36)=9.54$, $p<.01$; $F2(1,88)=7.96$, $P<.01$; last fixation duration: 243ms vs. 229ms: $F1(1,36)=7.87$, $p<.01$; $F2(1,88)=6.83$, $P<.01$; single fixation duration: 246ms vs. 229ms: $F1(1,36)=10.40$, $p<.01$; $F2(1,88)=8.66$, $p<.01$; gaze duration: 125ms vs. 98ms: $F1(1,36)=45.71$, $P<.001$; $F2(1,88)=36.13$. $p<.001$; go-past time: 140ms vs. 110ms: $F1(1,36)=29.74$, $p<.001$; $F2(1,88)=27.39$. $p<.001$; first pass re-reading time: both $Fs<1$).

Alternative high frequency words were also associated with later first landing positions than determiners (1.46 vs. 1.29 character spaces: $F1(1,36)=6.07$, $p<.05$; $F2(1,88)=14.13$, $p<.001$), and were less likely to be skipped (50% vs. 59%: $F1(1,36)=28.16$, $p<.001$; $F2(1,88)=30.71$, $p<.001$). Overall, this pattern replicates the finding that determiners are processed with relative ease compared to alternative high frequency words of the same length (e.g., Drieghe et al, 2008).

5.4.2.1.3. Word N+2: Readers engaged a more cautious reading style following an alternative high frequency word, with earlier first landing positions following this class of word compared to a determiner (2.46 vs. 2.89 character spaces: $F1(1,36)=28.67$, $p<.001$; $F2(1,88)=30.85$, $p<.001$). Word n+2 skipping was, however unaffected by word n+1 type (both $Fs<1$), as were all durational measures (gaze duration: F1(1,36)=2.50, p=.12; F2(1,88)=2.21, p=.14; all other Fs<1)

5.4.2.1.4. Spillover Region: A tendency for increased go-past time, driven in part by increased re-reading time, was present when word n+1 was a determiner, although this tendency was clearly not typical for all items (go-past time: 549ms vs. 521ms: F1(1,36)=16.26, p<.001; F2(1,88)=1.30, p=.26; first pass re-reading time: 71ms vs. 52ms: F1(1,36)=5.10, p<.05; F2(1,88)=2.82, p=.09). There were no other effects of word n+1 type in this region (first fixation duration: both Fs<1; gaze duration: F1(1,36)=1.63, p=.21; F2<1; first landing position: F1(1,36)=2.92, p=.09; F2(1,88)=2.48, p=.12).

5.4.2.2. Effects of Word N+2 Preview

5.4.2.2.1. Word N: As can be seen from Table 5.9, there was no evidence for a parafoveal-on-foveal effect stemming from word n+2 (single fixation duration: F1(3,108)=1.39, p=.25, F2<1; last fixation duration: F1(3,108)=1.65, p=.18; F2<1; all other Fs<1.2). This result is in agreement with Experiment 5 and suggests that the word n+2 plausibility manipulation was too remote to be detected by the reader.

Table 5.9. *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for Word N.*

	Identical	**Plausible**	**Anomalous**	**Illegal**
First Fix	247	252	248	248
Last Fix	246	249	241	246
Single Fix	250	257	250	252
Gaze	250	256	256	251
Go-Past	283	285	284	282
Re-Reading	33	29	28	31
Skip Prob	11	12	11	12
Landing	2.76	2.92	2.77	2.85

5.4.2.2.2. Word N+1: The word n+2 preview manipulation did not influence where within word n+1 first fixations fell (both Fs<1.1). There was also no evidence that it affected any of the individual fixation duration measures (first fixation duration: F1(3,108)=1.92, p=.13; F2<1; single fixation duration: F1(3,108)=1.50, p=.22; F2<1); last fixation duration: F1(3,108)=1.60, p=.19; F2<1), or the probability of regressing from word n+1 (both Fs<1).

There was however, a clear main effect of word n+2 preview on the probability of skipping word n+1 (F1(3,108)=4.06, p<.01; F1(3,264)=4.50, p<.01). As can be seen from Table 5.10, a lower skipping probability was associated with identical previews compared with any other preview type (identical vs. plausible: $F1(1,36)=10.77$, $p<.01$; $F2(1,92)=11.12$, $p<.01$; identical vs. anomalous: $F1(1,36)=6.23$, $p<.05$; $F2(1,92)=7.82$, $p<.01$; identical vs. illegal: $F1(1,36)=5.43$, $p<.05$; $F2(1,92)=5.63$, $p<.05$) [32]. Importantly, there was clearly no difference in skipping probability between the plausible and anomalous preview conditions (both Fs<1), suggesting this effect cannot be attributed to the extraction of word n+2 plausibility while fixating word n. This result is, however, difficult to interpret since a decision to skip word n+1 must have been made while fixating word n, at which point, both the identical and plausible previews should have been equally plausible and should therefore not have differentially affected word n+1 skipping rates.

Table 5.10. *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for Word N+1.*

	Identical	Plausible	Anomalous	Illegal
First Fix	240	231	235	240
Last Fix	241	231	235	238
Single Fix	241	232	236	241
Gaze	122	102	109	111
Go-Past	137	115	124	125
Re-Reading	15	12	15	13
Skip Prob	49	56	55	55
Landing	1.39	1.35	1.41	1.34

[32] All pairwise comparisons reported in this Results Section are collapsed over word n+1 type.

To determine whether this unexpected difference between identical and plausible conditions was the result of differences in predictability between these two conditions, a logistic regression was carried out for this subset of the data (identical and plausible conditions only) with condition treated as a dichotomous variable and predictability a continuous predictor. This analysis also demonstrated a significant difference between the identical and plausible conditions ($z=3.26$; $p<.01$), but with no effect of predictability ($z=1.02$; $p=.31$) and with no interaction between these two predictors ($z=-0.84$; $p=.40$)[33]. This result is unsurprising given the low predictability of word n+2 in the present study (<0.06) and strongly suggests that the predictability of word n+2 can be discounted as a possible explanation for this result.

As will be recalled from the Method section, there was a small but reliable difference in plausibility rating for the identical and plausible conditions. So to determine whether this could be driving the effect, another logistic regression was carried out on this dataset entering individual item plausibility ratings into the model as a continuous predictor variable. In contrast to when plausibility condition was considered as a categorical variable, there was clearly no effect of word n+2 plausibility when it was

[33] A maximal random effects structure was adopted for both subjects and items (Barr et al, 2013). This included random intercepts, and random slopes for word n+2 preview (identical vs. plausible), predictability (centred) and their interaction. Correlations were however omitted due to problems with convergence.

entered as a continuous predictor ($z=0.18$; $p=.86$)[34]. It seems therefore, that the small difference in plausibility between these two conditions cannot be driving the effect.

The remaining possibility is that it might have been driven by differences in the initial bigram or trigram frequencies of the identical and plausible previews. Indeed, in a study carried out by Pynte et al (2004), these sublexical properties were found to be responsible for differential skipping rates on a preceding word. This explanation seems somewhat unlikely in the present experiment, since any difference in the initial bi/trigram frequencies of the identical and plausible words should be negligible compared to the irregularity in the illegal nonword previews. If sublexical properties were driving this effect, one would expect that the illegal letter strings should have provided the divergent result – not identical words, as seen here. It seems therefore that this difference in skipping between the identical and plausible conditions is most likely the result of a type I error, with something about a subset of the words assigned to the identical condition exerting an inhibitory effect.

Both of the cumulative measures also revealed a main effect of the word n+2 preview manipulation (gaze duration: $F1(3,108)=4.45$, $P<.01$; $F2(3,264)=36.13$. $p<.001$; and go past time: $F1(3,108)=3.17$, $p<.05$;

[34] This model included a full random effects structure for both subjects and items including both random intercepts and slopes and correlations.

$F2(3,264)=3.69$, $p<.05$). However, since both of these effects disappeared when zeroes (skips) were removed from the calculation (gaze duration: both Fs<1; go-past time: $F1(3,108)=1.88$; $p=.14$; $F2<1$) this suggests that it was principally driven by skips. Therefore, aside from the effect seen in skipping rates, there was no evidence that word n+2 preview modulated word n+1 inspection.

5.4.2.2.3. Word N+2: As can be seen in Table 5.11, first landing position within word n+2 was unaffected by preview (both Fs<1), as was the probability of it being skipped ($F1<1$; $F2(3,264)=1.25$, $p=.29$).

There was, however, some evidence for an effect of preview in both first and single fixation duration, although statistically, this only approached significance in the latter (first fixation duration: $F1(3,108)=2.21$, $p=.09$; $F2(3,264)=1.97$, $p=.12$; single fixation duration: $F1(3,108)=2.38$, $p=.07$; $F2(3,264)=3.23$, $P<.05$). Pairwise comparisons revealed that, compared with when word n+2 had received an identical preview, neither the plausible nor illegal letter string previews significantly affected single fixation durations ($F1<1$; $F2(1,92)=3.54$, $p=.06$ and both Fs<1, respectively), single fixation duration was, however, reliably longer following an anomalous preview ($F1(1,36)=3.94$, $p=.05$, $F2(1,92)=6.49$, $p<.05$).

Table 5.11. *Fixation Time Measures (ms), Skipping Probabilities (%) and First Landing Positions (character spaces) for Word N+2.*

	Identical	Plausible	Anomalous	Illegal
First Fix	260	262	268	258
Last Fix	250	248	255	248
Single Fix	261	265	270	259
Gaze	286	309	298	297
Go-Past	342	355	355	357
Re-Reading	56	47	58	60
Skip Prob	6	5	7	5
Landing	2.71	2.66	2.64	2.68

The critical test to establish whether parafoveal plausibility had been extracted – a comparison between plausible and anomalous previews – showed no reliable difference ($F1(1,36)=1.25$, $p=.27$; $F2<1$). Nonetheless, the means here do replicate the longer inspection times seen with anomalous rather than plausible previews in both Experiments 4 and 5.

In contrast, as Table 5.11 suggests, the main effect in gaze duration ($F1(1,108)=3.64$, $p<.05$; $F2(3,264)=4.24$, $p<.01$), appears to have been orthographic in nature. Here, the general pattern reflects an increase in inspection time following all invalid previews, but no evidence to suggest that anomalous previews were associated with a higher level of interference.

Statistically, only the plausible previews resulted in increased gaze duration once word n+2 was fixated ($F1(1,36)=9.42$, $p<.01$; $F2(1,92)=20.05$, $p<.001$), with numerically and statistically weaker increases following the anomalous and nonword previews ($F1(1,36)=3.24$, $p=.08$; $F2(1,92)=3.78$, $p=.05$ and $F1(1,36)=2.50$, $p=.12$; $F2(1,92)=3.23$, $p=.07$, respectively). It appears, therefore, that once refixations are taken into account, the familiar pattern of increased inspection time following anomalous previews compared to the plausible condition disappeared. Indeed, the means here fall in the opposite direction, but not significantly so ($F1(1,36)=2.43$, $p=.12$; $F2(1,92)=2.51$, $p=.11$).

Any effects were clearly short-lived, since neither persisted into last fixation duration ($F1(3,108)=1.21$, $p=.31$; $F2<1$), go-past time (both $Fs<1.1$) or first pass re-reading ($F1(3,108)=1.17$, $p=.32$; $F2<1$); nor into the spillover region (see below).

5.4.2.2.4. Spillover Region: As Table 5.12. indicates, there were no reliable effects of the word n+2 preview manipulation in any measure within the spillover region (go-past time: $F1(3,108)=1.58$, $p=.20$; $F2(3,264)=1.29$, $p=.28$; all other $Fs<1$).

Table 5.12. *Fixation Time Measures (ms) and First Landing Positions (character spaces) for the Spillover Region*

	Identical	Plausible	Anomalous	Illegal
First Fix	249	253	248	247
Gaze	280	269	277	270
Go-Past	541	535	543	522
Re-Reading	61	66	66	52
Landing	4.47	4.53	4.53	4.65

5.4.3. General Discussion of Experiment 6

The design of the items in this experiment should have allowed word n+2 plausibility-related preview effects to optimally occur. However, despite finding that the anomalous previews caused the most disruption on word n+2 in single fixation duration, there was no difference between the plausible and anomalous conditions and no evidence for a plausibility-related preview effect in any other word n+2 measure. These findings (or lack thereof) leave open the possibility that the word n+1 plausibility-related preview effects reported in Experiment 4 were the result of either a mislocated fixation or a quick succession of lexical processing, both of which are consistent with the E-Z Reader model of eye movement control.

There was also no evidence that the word n+2 plausibility manipulation influenced fixations on either word n or n+1. The absence of a plausibility-related parafoveal-on-foveal effect replicates Experiment 5 and further suggests that these effects do not survive the presence of an intervening word. The absence of a delayed parafoveal-on-foveal effect does, however, contrast with Experiment 5, where an effect in first landing position was found. It also contrasts with Experiments 2 and 3 where word n+2 preview appeared to modulate word n+1 skipping rates. As previously discussed, delayed parafoveal-on-foveal effects are far from stable across experiments. Some labratories report obtaining such effects in durational measures (e.g., Kliegl et al 2007; Risse & Kliegl, 2011), while others report obtaining them only in saccadic measures (e.g. Angele & Rayner, 2011; Pynte et al, 2004).

Since these effects were not present here with the anomalous (i.e., real word) previews, this leaves open the possibility that such effects could be driven by a low level attentional scan that detects upcoming irregularity (e.g., Angele & Rayner, 2011). It has also been suggested that delayed parafoveal-on-foveal effects could result from a failed skip of word n+1 followed by a stay and process response (e.g., Schotter et al, 2012). On a theoretical level, such an explanation does not appear to explain the variety of results across experiments, such as their presence and/or direction. This explanation will, however, be fully explored in Chapter 6, where the oculomotor response to a mislocated fixation is experimentally tested.

The only evidence for a word n+2 preview effect was orthographic in nature. This result adds to several recently published studies in showing the existence of an orthographic word n+2 preview effect (e.g., Angele & Rayner, 2011; Experiment 3 of this thesis; Radach et al, 2007, 2013). But, since the effect was present when word n+2 had contained real-word previews, this removes the possibility that the low level attentional scan can account for the effect. However, a recently published simulation with the E-Z Reader model shows that word n+2 preview effects can be accounted for by the model via a double attention shift, provided they are orthographic in nature: the precise pattern reported here.

Taken together, aside from the numerical trend towards a plausibility-related preview effect on word n+2, there was no other evidence to suggest that word n+2 plausibility was extracted while fixating word n. One problem with all boundary paradigm experiments that manipulate word n+2 previews, however, is that evidence for word n+2 effects will become diluted if word n+1 is also fixated. This is because, upon fixation of word n+1, word n+2 will always be available in its appropriate form for parafoveal processing. This may, in turn, dilute any preview effects that result from an inaccurate preview while inspecting word n – especially since word n+2 preview effects are likely to be smaller in magnitude than those stemming from an immediately adjacent word. Thus, the present set of results leave open the possibility that the plausibility-related preview effects reported in Experiment 4 might have resulted from either a quick succession of lexical processing, or a mislocated fixation. The absence of clear effects cannot, however, be taken as evidence that either of these processes were necessarily driving the plausibility-related preview effects reported in Experiment 4, just that they might be.

5.5. General Discussion of Chapter 5

The experiments presented in this Chapter provide clear evidence for plausibility-related preview effects originating from word n+1, but not from word n+2. The trends in both Experiments 5 and 6 were not consistent and unreliable. Thus, while the results here provide, like others before (e.g., Kennedy et al, 2004; Murray, 1998; Murray, 2006; Murray & Rowan, 1998; Starr & Inhoff, 2004), further evidence for the parafoveal extraction of plausibility-related information, the possibility that these effects might have been driven by a fast succession of lexical processing, or a mislocated fixation coupled with a stay and process response, cannot be discounted.

The plausibility-related preview effects in Experiment 4 cast some doubt over whether the depth of orthography hypothesis (Schotter, 2013) can provide a plausible account for why previous semantic preview benefit experiments have produced discrepant findings. If semantic preview benefit occurs in German (e.g., Hohenstein et al, 2010; Hohenstein & Kliegl, 2013) but not in English (e.g., Rayner et al 1986; 2014; Schotter, 2013) solely because languages with deeper orthographies (such as English) restrict the time available for the extraction of parafoveal meaning, then the plausibility-related effects of Experiment 4 should not have occurred. These effects suggest that there was time to extract parafoveal meaning and, further, to perform some degree of high-level processing, in a language considered to have a deep orthography.

The word n+1 plausibility-related trends and effects reported in this chapter suggest that different previews elicit differing degrees of interference. This fits with the way some researchers suggest a need to reconceptualise the term "preview benefit" to incorporate not just potential benefits, but also costs resulting from a word change (Murray, Rayner & Wakeford, 2013; Risse & Kliegl, 2013). As discussed above, it appears that meaning-based interference could potentially provide some explanation for the lack of preview benefit seen in many studies, with the degree of correspondence between semantic associate meanings being too divergent to provide any facilitation (e.g., Rayner et al, 1986; 2014). Equally, positive evidence could be attributed to a closer correspondence in meaning (e.g., Hohenstien & Kliegl, 2013; Schotter, 2013)

Similarly, positive evidence for semantic preview benefit might also be driven by differences between the plausibility of semantically related and unrelated previews. While such an explanation appears compatible with White et al's (2008) finding of a semantic preview benefit on the second constituent of a two-part compound word (see Experiment 4 General Discussion), this interpretation is based on the only example item that they provided.

A recently published paper by Rayner and Schotter (2014) provides a more concrete example of how differences in plausibility might mimic semantic preview effects. They report obtaining no semantic preview benefit with uncapitalised words, but obtained a semantic preview benefit when the initial letters of previews and targets were capitalised. Whether, as the authors suggest, this result reflects faster orthographic processing due to the capitalisation, which in turn allows more time for semantic pre-processing of the upcoming word, or whether it simply encourages an atypical distribution of attention to the parafoveal word, cannot be determined from their study. Inspection of their materials (a full list of which is appended to their paper) does, however, demonstrate that the previews defined as semantically unrelated to the target word tended to be either severely implausible or anomalous, while those that were defined as semantically related were typically plausible. While there were some exceptions to this rule, this systematic variation in the items is quite striking. Thus, these recently published results fit nicely with the hypothesis that semantic preview effects might actually be caused by differing levels of interference originating from higher level processing.

The extent to which orthographic regularity, salience and semantic interference contribute to variation in the pattern of results obtained in parafoveal 'priming' experiments remains an open question, but what is now clear is that semantic information can be extracted from a parafoveal word, and that this information can be employed in higher-level sentence processing before the eye moves onto that word.

The suggestion that plausibility-related preview effects reflect the engagement of higher-level processing on a parafoveal word might be challenged by some who suggest that the detection of anomaly can reflect a lexical, rather than a post-lexical effect.

The anomalous words here comprised both syntactic and semantic anomalies. The former were cases where the word fell into the wrong syntactic category, for example, a verb followed by a verb. The latter occurred with words of the correct form class, but where these violated possible selectional restrictions, for example: "*fed the chair*" (only living things require sustenance). It has previously been suggested (e.g., West & Stanovich, 1986; Wright & Garrett, 1984) that syntactic category effects can reflect pre-lexical selection rather than post-lexical integration processes. A similar argument might be made regarding selectional restrictions. Thus, these effects might have been driven by pre-lexical rather than advanced post-lexical processing.

However, the proposal that syntactic restrictions might exert a pre-lexical effect is contentious. While there is a wealth of research supporting a role for, for example, frequency and phonology exerting an effect during pre-lexical word processing (e.g., Howes & Solomon, 1951; Whaley, 1978; Rubenstein, Lewis & Rubenstein, 1971), there is little evidence to suggest that syntactic restrictions play an early role. Aside from the general argument that to know if a word fits syntactically, you must first know what that word is, a recent eye tracking study also indicates that syntactic categories are not involved in the process of lexical identification. In their study, Scougal and Murray (2009) obtained no evidence to suggest that target word frequency (i.e., a pre-lexical effect) interacted with whether or not the target word was syntactically acceptable. If both factors were indeed pre-lexical, one would expect evidence of an interaction between them, since, for example, a restriction in target set size, based on syntactic category, would be expected to influence the magnitude of the frequency effect. It seems unlikely therefore on both logical and empirical grounds that the effects of syntactic anomaly could be accounted for by a pre-lexical process.

Turning to selectional restrictions, Warren and McConnell (2007) have recently re-introduced the possibility that information relating to these might be involved in the lexical recognition process; leaving open the possibility that such violations could potentially exert a pre- rather than post-lexical effect. This proposal, which dates back to Katz and Fodor (1963), might suggest that the observed difference between plausible and anomalous word previews could have been driven by pre-lexical rather than post-lexical pre-processing.

Again, however, there is little evidence for such a proposal. It assumes that the meaning of a word can influence the word recognition process before the word's meaning is known, presumably by some pre-selection of a set of 'compatible' candidate lexical entries, but there seems to be no evidence for such a selection effect. Furthermore, Schotter (2013) has recently provided evidence (via post hoc LMM analyses) that seems to contradict the proposal. As will be recalled, Schotter obtained a semantic preview benefit for synonyms but not for semantically associated word pairs. To determine whether the lack of difference between the semantically related and unrelated words might have been caused by syntactic or semantic anomalies within these two conditions, she ran a series of sub analyses. In addition to finding evidence that both types of anomaly behaved similarly, she also reported that both types of anomaly influenced later measures (i.e., those involving regressions), which she interpreted as both types of anomaly being implicated in post lexical integrative processes[35]. It therefore seems most likely that the results of Experiment 4 implicate higher-level, rather than pre-lexical processing.

In a similar vein, to account for the early time course of severe implausibility/anomaly on the eye movement record, Staub et al (2007) make

[35] Indeed, this is similar to the pattern of results here, where the plausibility-related preview benefits only achieved significance in go-past time.

the case that severe implausibility, such as anomaly, may be detected via an anticipation-led disconfirmation of an expected category (e.g., encountering a non-animate object when an animate object had been expected). Such an explanation does, however, raise several questions, not least of which is how it can be known that a word does not fit an expected category prior to the identity of that word being known.

5.5.1. Implications of Plausibility-Related Preview Effects for Models of Eye Movement Control.

SWIFT: The current version of the SWIFT model is clearly too underspecified to account for the present set of results. While at a theoretical level, the model should be able to account for plausibility-related preview benefits, until such a mechanism is incorporated within the model, it is unknown whether these effects could be successfully modelled within SWIFT's parallel lexical processing architecture. It should be added that the incorporation of such a mechanism must also not be to the detriment of the model's current capabilities in accounting for existing benchmark effects. It seems likely that adding such a mechanism to the SWIFT model would greatly increase its complexity, since there would be unfolding lexical and higher-level processing effects occurring simultaneously and this might well pose a significant modelling challenge for the architects of SWIFT.

E-Z Reader: As discussed earlier in this chapter, Schotter et al (2014) have provided a series of simulations demonstrating that the E-Z Reader model can account for semantic preview benefit. It will also be recalled, however, that questions were raised regarding the methodology of these simulations. It was argued that the measures of whether the L2 stage of word n+1 had been reached, and time spent within that stage, do not necessarily align with the suggestion that the E-Z Reader model can account for semantic preview benefit. These simulations can only inform on that question once the analyses are restricted to cases where the L2 stage had been reached *after* the saccade to word n+1 had been committed to action. Any cases occurring before this should be excluded, since according to the model, these words will be skipped preventing them from contributing to a semantic preview effect on word n+1.

At a theoretical level, it is also difficult to conceptualise how a semantic preview benefit could exist within the E-Z Reader model's framework. According to this model, if some degree of semantic preprocessing can occur on a word prior to its fixation, then upon fixation, the subsequent change in word should surely cause some level of interference, whether the word is a semantic associate (e.g., from "coin" to "bank"; taken from Rayner & Schotter, 2014) or not. Rayner and Schotter (2014) acknowledge this conundrum but suggest that the word could be sufficiently processed by the time that it is fixated such that the reader might not register the change in word, which they suggest could account for the lack of disruption. Such an explanation, however, appears to necessitate the suggestion that word n+1 had been fully lexically processed by the time the word was fixated, otherwise attention should still be directed to word n+1 and the change detected. Furthermore, while such an explanation might account for the lack of disruption observed on a given target word, the processing of an incorrect semantic associate seems likely to be detected further downstream, presumably resulting in a tendency to make a regression back into the target word once it has been passed (e.g., Frazier & Rayner, 1982; Kennedy & Murray, 1987). To fully explore this explanation, it will be necessary to elicit a semantic preview effect using semantic associates designed to cause the parser processing difficulties later downstream, and then measuring whether these instances result in an increased tendency to refixate the target word.

But regardless of whether the E-Z Reader model is capable of accounting for semantic preview benefit, it is unlikely that it could accommodate the loose coupling between fixation location and attention needed to account for a plausibility-related preview effect. This would necessitate not only the semantic code of a parafoveal word being extracted prior to fixation, but also that that information could be utilised at a high-level as well.

While the E-Z Reader model has recently incorporated a higher level processing module within its architecture, given the simulations presented by

Schotter et al (2014), it seems highly unlikely that lexical processing of the parafoveal word could routinely progress far enough to permit the type of plausibility-related preview effects reported here. It therefore appears that a quick succession of lexical processing is an unlikely explanation for the present results. The remaining possibility is that these effects could be accounted for by a mislocated fixation followed by a stay and process response; the plausibility of such a response to a mislocated fixation is explored further in the following chapter.

CHAPTER 6

The Consequences of Mislocated Fixations During Reading

6.1. Introduction

A recurrent theme throughout this thesis has been the notion that many effects suggestive of parallel lexical processing can be explained within a serial framework providing one concedes that many mislocated fixations will be followed by a stay and process response. This concept was first popularised by Dreighe et al (2008), and as discussed in Chapters 4 and 5, it can be invoked to explain evidence pertaining to parafoveal meaning effects (e.g., Experiment 4 of this thesis; Hohenstein & Kliegl, 2013; Schotter, 2013), lexical parafoveal-on-foveal effects (e.g., Hyönä & Bertram, 2004; Inhoff et al, 2000; Kennedy, 1998, 2000; Kennedy & Pynte, 2005; Kliegl et al 2006; 2007), and so-called delayed parafoveal-on-foveal effects (e.g., Angele & Rayner, 2011; Kliegl et al, 2007; Risse et al, 2013). The concept of a mislocated fixation coupled with a stay and process response has thus played a central role in maintaining the viability of the E-Z Reader model since the above-mentioned effects filtered into the literature. To date, researchers have been unable to assess the merit of this auxiliary assumption, since it has been argued that it is impossible to know whether a fixation is mislocated, and therefore, that the response it typically induces is difficult to establish (Drieghe, 2011). As a consequence, the E-Z Reader model has lost its ability to make transparent predictions, at least as far as parafoveal meaning effects and parafoveal-on-foveal effects are concerned, causing it to “take on a ghostly quality” – the precise criticism proponents of the E-Z Reader model exact against parallel models of eye movement control (Reichle et al, 2009; p116). Given the central role that this concept now plays in the explanatory adequacy of serial models, the following experiment sought to experimentally test whether there is any foundation to the mislocated fixations hypothesis as advocated by Drieghe et al (2008) by artificially inducing mislocated fixations using a text shift paradigm.

McConkie et al coined the term ‘mislocated fixation’ in 1988 to refer to fixations that, they hypothesised, had missed their intended target word, and in so doing, erroneously landed on an adjacent word. As will be recalled from Chapter 1, these conclusions were drawn after McConkie et al noticed a series of systematic variations in the landing site distributions of first fixations within words. First, their landing site distributions assumed a normal distribution with a peak offset to the left of word centre. Second, there was a left- or rightward shift in that peak for saccades that travel more or less than 7-character spaces, respectively. McConkie et al interpreted this variation as reflecting fixations that had under- or overshot the preferred central location, with the magnitude of this error being approximately half a character space for every character space the saccade deviated from 7-letters. McConkie et al referred to this error component as a systematic range error. Third and finally, they reported a random error component, which also assumed a normal distribution around the word’s centre and was moderated by launch site fixation duration with shorter durations inducing flatter landing site distributions and therefore a wider spread of error. The principles of oculomotor error outlined in McConkie et al’s seminal paper have proved incredibly influential for models of eye movement control, with both the E-Z Reader (Reichle et al, 2009) and SWIFT (Schad & Engbert, 2013) models incorporating their principles into their respective architectures.

McConkie et al’s finding that readers typically fixate just to the left of word centre aligns with the results of earlier research (e.g., O’Regan, 1981; Rayner, 1979) and is a position commonly referred to as the ‘preferred viewing location’ (PVL). As briefly discussed in Chapter 1, it makes sense that readers should aim for this location since it aligns, approximately at least, with the location where words are identified most efficiently when presented in isolation; a location known as the ‘optimal viewing position’ (OVP; O’Regan, 1981; O’Regan & Jacobs, 1992; O'Regan, Lévy-Schoen, Pynte & Brugaillère).

One finding that, until recently, proved difficult to reconcile with this pattern of effects was the inverted optimal viewing position (IOVP) effect first reported by Vitu et al (2001). They found that when first fixation durations were plotted as a function of within-word landing site, the distribution, although still approximately normal, showed the longest durations at word centre, decreasing with increasing eccentricities. This pattern thus appeared to show an advantage to fixating the peripheral regions of a word, which according to the OVP research, is the least optimal place for efficient word identification.

To account for these apparently contradictory IOVP effects, Nuthmann et al (2005; 2007) drew on McConkie et al's (1988) insights about mislocated fixations. They hypothesised that the short first fixation durations near word boundaries may be driven by fixations missing their intended target words followed by a fast error-correcting saccade. This account explains why first fixation durations are shortest at word-eccentric positions, since here, mislocated fixations aimed at the immediately adjacent words will be more prevalent than at word-centric positions. Indeed, McConkie, Kerr, Reddix, Zola and Jacobs' (1989) report that refixations are most prevalent near word boundaries, indicating that these positions are not as 'optimal' as the short first fixation durations in the IOVP distributions might otherwise suggest. Nuthmann et al (2007) ran a numerical simulation with the SWIFT model (Version 2; Kliegl et al, 2005) to assess the likely frequency of mislocated fixations during reading. These simulations estimated that as many as 23.2% of all fixations might be 'mislocated', that is, they missed their intended target word.

The simulations conducted by Nuthmann et al clearly indicate that mislocated fixations are unlikely to be a rare phenomenon and as such, any systematic response to their occurrence could have a significant impact on the eye movement record. Critically, unlike Nuthmann et al who suggest that the response to a mislocated fixation is to program a quick error correcting saccade, in order to explain lexical parafoveal-on-foveal effects, Rayner et al (2004) suggest that the opposite occurs; specifically, that mislocated fixations will frequently be followed by a stay and process response. This permits a degree of decoupling between fixation location and the locus of attention, which in turn allows for the lexical properties of a parafoveal word to be extracted and/or expressed on the preceding word – all within a serial framework[36]. This hypothesis necessarily restricts mislocated fixations' contribution to, for example, in the case of parafoveal-on-foveal effects, to progressive saccades that undershoot their targets. Nuthmann et al (2007) estimate that such fixations might comprise as much as 13% of all progressive fixations and therefore potentially enough to be driving this sort of effect.

Proponents of the E-Z Reader model have been quick to adopt the possibility of a stay and process response to a mislocated fixation to account for parafoveal-on-foveal effects (e.g., Schotter et al, 2012). They have, however, remained silent regarding the observation made by Kliegl and Engbert (2011) that this response is incongruent with the fast error correcting response advocated by Nuthmann et al (2005; 2007) and the account that this provides for the IOVP effect.

[36] Note that this hypothesis only permits orthodox parafoveal-on-foveal effects since the effect should imitate the processing that *would have* taken place *if* the parafoveal word had actually been fixated. This account is therefore unable to account for unorthodox parafoveal-on-foveal effects.

Drieghe et al (2008) suggest that they have obtained empirical support for the mislocated fixations account of parafoveal-on-foveal effects. Using the boundary paradigm, they presented participants with sentences containing two critical words: a five-letter high or low frequency noun (word n; e.g., “child”) followed by another word that was five or more characters long (word n+1; e.g., “performing”) that received either an identical or illegal (e.g., “pxvforming”) preview prior to fixation. They analysed fixation durations on word n as a function of word n+1 preview. They predicted that if parafoveal-on-foveal effects were caused by mislocated fixations, then (a) foveal word frequency should not modulate any parafoveal-on-foveal effect, (b) parafoveal-on-foveal effects should be confined to the final characters of the foveal word, and (c) since McConkie et al (1988) reported saccadic undershoots were most prevalent following long saccades (i.e., due to systematic error), then any parafoveal-on-foveal effect observed should be correlated with incoming saccade length. Drieghe et al report that these three hypotheses were confirmed. Indeed, the only evidence for a parafoveal-on-foveal effect in single fixation duration was on the final letter of word n, where there was an 87ms increase when the parafoveal word contained the illegal preview. There was no hint of a parafoveal-on-foveal effect on the duration of fixations falling on any other preceding letter. This research therefore appears to support the stay and process response to mislocated fixations.

Kennedy (2008), however, has shown that parafoveal-on-foveal effects are not necessarily confined to the final characters of the foveal word. Using linear mixed effects modelling on the Dundee Corpus, he found that parafoveal familiarity[37] influenced single fixation durations on the foveal word and, importantly, this effect did not interact with the length of the incoming saccade. Additionally, the effect of familiarity survived from letters -2 to -7 of the foveal word. These results lie in stark contrast to those reported by Drieghe et al (2008) and would appear to be in contrast with a mislocated fixations account of parafoveal-on-foveal effects. Indeed, Kennedy further suggests that Drieghe et al's results are also incongruent with the mislocated fixations view of parafoveal-on-foveal effects, since such an account should predict a more graded parafoveal-on-foveal effect on the foveal word, not one that is confined to the final letter. Proponents of the E-Z Reader model have nevertheless remained steadfast in advocating the stay and process response to a mislocated fixation and its potential in explaining lexical parafoveal-on-foveal effects (e.g., Schotter et al, 2012).

It is clear that a great deal of theoretical importance has been placed upon whether mislocated fixations can account for parafoveal-on-foveal effects. Unfortunately, as Drieghe (2011) states "…the problem with the mislocated fixations account is, of course, that there is no way to

[37] Kennedy (2008) refers to lexical familiarity as the "cumulative lexical frequency" of a word, specifically "…for each word, the summed frequency of all words sharing that token's initial trigram." (p. 6)

experimentally determine whether a saccade has been mislocated or not" (p. 848). Drieghe's observation that mislocated fixations cannot be experimentally determined may, however, have been premature; one paradigm that appears to lend itself quite naturally to such an investigation is the text shift paradigm (O'Regan, 1981). This paradigm operates in very much the same way as the standard boundary paradigm, except that upon passing an invisible boundary, the sentence either remains static, shifts to the left or shifts to the right. Similarly to the standard boundary paradigm, saccadic suppression (Matin, 1974) renders most readers oblivious to these manipulations, permitting observation of how the oculomotor system responds to these artificially induced mislocated fixations.

Feng (2009, Experiment 3) used the text shift paradigm to measure reader's responses to a series of artificially induced mislocated fixations. He asked participants to read stories on a screen; during the first five pages the text remained static, thereafter the page of text either (a) remained static, (b) shifted to the left or (c) shifted to the right by 1 to 3 character spaces on every 8^{th} to 12^{th} saccade. Feng reported clear IOVP functions for words based on actual (i.e., post shift) landing sites, but no clear IOVP functions when the targeted word's landing sites were plotted as a function of launch site (i.e., pre shift) positions of those words. Feng took this as evidence that – providing one assumes the IOVP function is driven by mislocated fixations as proposed by Nuthmann et al (2005) – the detection of the mislocated fixation must have been based on retinal input rather than an efference copy written before the eye arrived at the erroneous fixation location.

More importantly for the present discussion, Feng also reported an asymmetry in response depending on whether the text shifted to the left or to the right. Following a left shift (i.e., a simulated overshoot), there was an increase in the number of regressions but no difference in first fixation duration compared with when the text remained static. Conversely, first fixation durations were markedly shorter following a rightward shift (i.e., a simulated undershoot) compared to the no shift condition, although the onward progression of the eyes was undisrupted. Together, these results suggest that mislocated fixations are detected via retinal input, but that the response differs depending on whether they are over- or under-shoots. Feng suggests that the left shift data do not support quick error correcting fixations as hypothesised by Nuthmann et al (2005); rather, he implicates cognitive processing in the response observed for the left shift data. Critically for the present discussion, however, the right shift data appear to support a quick relocation strategy[38]. Specifically, there was no support for longer durations, as one might expect if a parafoveal word was being processed from an erroneous location. Rather, fixations were shorter, as would be predicted if the decision is to relocate the fixation.

[38] Although not quick enough to be based on efference copy as hypothesized by Nuthmann et al (2005).

While these results appear to be inconsistent with Drieghe et al's hypothesis regarding the likely response to mislocated fixations, a stronger test is possible. In the present study, participants were presented with sentences in which a critical word was either a high or low frequency word. An invisible boundary was located immediately prior to this word, and, upon being passed, the entire sentence either (a) remained static, (b) shifted 2-character spaces to the left, or (c) shifted 2-character spaces to the right. A sentence shift of two-character spaces was used since this should be large enough to generate a reasonable proportion of mislocated fixations without attracting conscious attention. Using this design, it is possible to assess whether lexical processing commences immediately after a shift. If so, a frequency effect should be present and this could be taken as support for a stay and process response to the mislocation. Alternatively, the absence of evidence for lexical processing following a sentence shift would appear to implicate a quick error correcting strategy – especially if accompanied by shorter first fixation durations within that region.

6.2. Method

6.2.1. Participants

Thirty native English speakers with normal or corrected-to-normal vision and with no known reading difficulties took part in the experiment. Each received course credit for their participation.

6.2.2. Materials and Design

Ninety-six experimental sentence frames were constructed. Each sentence contained a two-word critical region comprising a 7 or 8 character high frequency (M=58/million) verb followed by either a high (M=166/million) or low (M=6/million) frequency 5 or 6 character noun. Word frequencies were estimated using the Kuçera and Francis (1967) norms. Sentence lengths ranged from 60 to 80 characters to enable them fall on a single line of the CRT display. Verbs were relatively long and high frequency with predictable endings (such as, "ing", "ed") to help ensure that they tended to capture a single fixation followed by a relatively long progressive saccade into the adjacent noun, thus creating conditions under which an undershoot of the noun might be considered likely.

Initial presentation position of each sentence was identical across all shift conditions; however, once the eye passed an invisible boundary located half way between the verb and noun, the sentence shifted either (a) 2 character spaces to the left, (b) 2 character spaces to the right, or (c) did not shift at all (the sentence was replaced by itself). As can be seen from the Figure below, this resulted in a 2 (noun frequency) x3 (shift condition) design with a total of 6 conditions:

High Frequency Noun

6a) Amanda had carefully prepared|tests for the final year students. (--)

6b) Amanda had carefully prepared tests for the final year students. (<-)

6c) Amanda had carefully prepared tests for the final year students. (->)

Low Frequency Noun

6d) Amanda had carefully prepared|robes for the final year students. (--)

6e) Amanda had carefully prepared robes for the final year students. (<-)

6f) Amanda had carefully prepared robes for the final year students. (->)

Figure 6.1. *An example item with either a high (6a-6c) or low (6d-6f) frequency word in the position of n+1 presented in each shift condition. Symbols (---), (<--) and (→) present relative sentence positions for the no shift, left shift and right shift conditions, respectively. The "|" symbol denotes the invisible boundary location.*

Six counterbalanced item files were constructed. Each participant experienced all preview conditions across an equal number of items, but saw only one version of each item.

To ensure normal reading for comprehension, 20 of the experimental items were followed by a comprehension question. In addition, a further 32 items of similar structure, but with no contingent shift, were constructed as filler items. Eight separate practice items, again with no contingent shifts, preceded the experimental items; half of these were accompanied by a comprehension question. Consequently, during the main experimental block, 25% of the items shifted left, 25% shifted right and 50% had no shift at all.

6.2.3. Apparatus

Participant’s eye movements were tracked using an EyeLink 2000 eye tracker (SR Reasearch Ltd, Kanata, Ontario, Canada). The tracker was set to tower mode with pupil position sampled every millisecond. Sentences were presented in black monopitch text on a white background using a ViewSonic Graphics Series G220fb monitor with refresh rate set at 120Hz. At a viewing distance of approximately 80cm, each character subtended approximately 0.32 degrees of a visual angle. Viewing was binocular but only the movement of one eye was recorded.

The contingent change was achieved by replacing a pre-shift .png image of the sentence with a post-shift .png image of the same sentence but in the relevant post-shift position. To ensure strict comparability between display conditions, the contingent change was also employed in the identical preview condition (i.e., the sentence was replaced by itself). Given the nature of the display change, extra care was exercised in order to maximise the chance that it would occur during saccadic suppression (Matin, 1974) with triggering of the change only occurring when saccadic velocity exceeded 20 degrees per second as the eye crossed the invisible boundary. Inefficient display changes were identified after testing had been completed and those deemed unsatisfactory (e.g., they occurred too late) were deleted prior to any subsequent analysis. The process by which these cases were identified is outlined fully in the Results and Discussion section below.

6.2.4. Procedure

The procedure was identical to Experiments 2 to 6 with the following exceptions. First, participants were not required to use a bite bar. Second, the pre-sentence fixation marker was in the form of a small dot, not a cross; if fixations on this dot were deemed unsatisfactory, the eyes were recalibrated prior to the next trial commencing, but if satisfactory, the eyes were only re-calibrated once every tenth trial. Third, there were two experimental sessions, in Block 1 participants were exposed to the 96 items for the first time, in Block 2, they received another exposure of these same items but in a different random order. The entire experiment lasted approximately 80 minutes and participants were given a break between the two blocks and whenever they desired. All other aspects of the experimental procedure remained the same as in the previous experiments using the Dr Bouis eye tracker (see Chapter 4, Section 4.2.1.4).

6.3. Results and Discussion

The example below illustrates the three zones on which separate analyses were conducted:

```
Zone                   1       2         3
Amanda had carefully|prepared|tests|for the final|year students.
```

Pre and post shift data were extracted separately. Pre-shift analyses were restricted to data on the verb prior to the shift taking place (zone 1).

Post-shift data contributed towards analyses on the noun and spillover regions (zones 2 & 3, respectively). These were re-aligned to match the changed letter positions following the shift (if any) in order to examine effects of shift and frequency within these regions.

A number of fixation time measures were computed for each of these zones: first, last and single fixation durations, first landing position and skipping probability; all of which are fully described in Chapter 4, section 4.2.2. Gaze and go-past were also calculated but unlike the previous two chapters, skipped words were not attributed a value of zero. This decision was based upon the knowledge that the sentence shift would occasionally create artifactual word skipping, meaning that no assumptions can be drawn regarding whether a skipped noun might have been processed while fixating adjacent words. Refixation probability was also calculated on the noun to determine the probability of it being refixated prior to a saccade exiting it in either direction during first pass reading.

For cumulative measures, fixations with durations below 80ms or above 1500ms were deleted; the upper threshold was reduced to 1000ms for all other fixation based measures. Trials in which the shift was not executed efficiently were also removed from the analyses, determined by (a) a shift being executed prior to the boundary having been crossed or (b) when the shift occurred more than 9ms into the onset of a post-shift fixation. Overall, 19% of the data were lost as a result of these measures. Finally, any participant who, after testing, reported noticing more than 6% of the sentence shifts was replaced. This was only necessary on 2 occasions, since the majority of participants were oblivious to these manipulations. Participants clearly read the sentences carefully and with relative ease, achieving 91% accuracy in the comprehension questions.

Skipping and refixation probabilities were analysed using logistic mixed effects regressions, all other measures were analysed using linear mixed effects models (LMMs) treating subjects and items as crossed random effects. These analyses were chosen as an alternative to the more traditional ANOVA as they are compatible with unbalanced datasets (Baayen, Davidson & Bates, 2008) and it was anticipated that many analyses here would likely involve significant amounts of missing data. All models were run within the R statistical programming environment (R: A language and environment for statistical computing, The Core Team, 2012), using the lmer function of the lme4 package (Bates, Maechler, Bolker, 2014). As a direct comparison between the two shift conditions was not of theoretical interest, separate models were built to compare (a) the no shift condition against the left shift condition, and (b) the no shift condition against the right shift condition; this resulted in each model comprising two fixed effects: noun frequency (high versus low frequency) and shift (shift versus no shift). Attempts were made to perform analyses with a maximal random effects structure (Barr, Levy, Scheepers & Tily, 2013), however, since this resulted in the models failing to converge, all correlations relating to the random effects were removed[39]. This step enabled all random effects structures for all models reported in this section to retain random intercepts and slopes for frequency and shift and a random slope for their interaction, both for items and subjects.

Since log transformed data produced the same critical results as the raw data, all analyses were – following what appears to be an emerging consensus in this field – conducted on the raw, untransformed dataset. Inferential statistics are reported as follows: for LMMs regression coefficients (*b*), standard errors (*SEs*) and *t*-values (*t*) are provided, a fixed effect is interpreted as significant if the *t*-value exceeds 1.96 (indicating alpha is <.05), marginally significant if it falls between 1.64 and 1.96 and not significant if it is below 1.64. For logistic mixed effects regressions, a *z*-value (*z*) is reported together with its associated significance level (*p*-value).

6.3.1. Verb

Left Shift: As the means presented in Table 6.1 indicate, inspection times, skipping probability and first landing position were all unaffected by the upcoming word's frequency class (first landing position *b*=0.08, *SE*=0.07, *t*=1.12; all remaining *t*s<1) and shift condition (last fixation duration: *b*=3.05, *SE*=3.03, *t*=1.00; first fixation duration: *b*=2.89, *SE*=2.82, *t*=1.02; first landing position *b*=-0.10, *SE*=0.07, *t*=-1.40; all other *t*s<1); nor was there any evidence

[39] Barr et al (2013) suggest that this step is preferable to removing random intercepts or slopes since LMMs lacking correlations perform similarly to those boasting a full maximal structure.

of a significant interaction between these two factors (go-past time: *b*=18.34, *SE*=13.29, *t*=1.38; all other *t*s<1).

Table 6.1. *Fixation Time Measures (ms) and Skipping Probabilities (%) and First Landing Positions (character Spaces) for the Verb.*

	No Shift		**Left Shift**		**Right Shift**	
	LF	**HF**	**LF**	**HF**	**LF**	**HF**
First Fix	225	225	227	227	227	227
Last Fix	221	220	223	222	226	226
Single Fix	230	228	232	232	233	234
Gaze	257	260	258	257	262	260
Go-Past	289	281	276	286	291	280
Skip Prob	8	6	7	7	8	8
Landing	3.15	3.19	3.02	3.14	3.09	3.08

Right Shift: Aside from a numerical trend towards longer go-past times when the word to the right was a low compared to a high frequency word (290ms vs. 281ms *b*=-8.08, *SE*=5.79, *t*=-1.52), no other measures revealed any evidence for a frequency driven parafoveal-on-foveal effect (all *t*s<1). Last fixation duration was marginally higher prior to a rightward shift, however, since this effect is numerically small (221ms vs. 226ms), statistically weak (*b*=5.09, *SE*=3.10, *t*=1.64) and confined to last fixation duration (single fixation duration: *b*=-3.55, *SE*=2.74, *t*=-1.30; first landing position: *b*=-0.09, *SE*=0.07, *t*=1.21; all other measures *t*s<1) and to the rightward-shift dataset alone (see above), it is probable that it is spurious, especially since it precedes the actual execution of the shift. No interactions between frequency and shift were present (all *t*s<1).

6.3.2. Noun

In order to plot landing site distributions, each noun was divided into 5 subregions referred to as quintiles. For 5-letter nouns, 1-quintile corresponds to 1-character space, for 6-letter nouns 1-quintile corresponds to 1.2 character spaces. Fixations falling on the spaces either side of the noun were not included in the following distributions.

The proportion of first fixations landing within quintiles 1 to 5 of the noun are presented below in Figure 6.2[40]. It is clear from Figure 6.2.A, that when the sentence remained static, the inverted U ('OVP') function, as obtained by McConkie et al (1988) was present. Unsurprisingly, the peak of this distribution was shifted to the right following a simulated overshoot (6.2.B) and shifted to the left following a simulated undershoot (6.2.C). As this selection of graphs indicates, the first landing position within the noun did not interact with its frequency. These statistics along with all other analyses are reported below with the left and right shift conditions considered separately.

[40] For clarity, these Figures include all fixations falling on the noun but exclude fixations that fell on the half character space located either side of the noun.

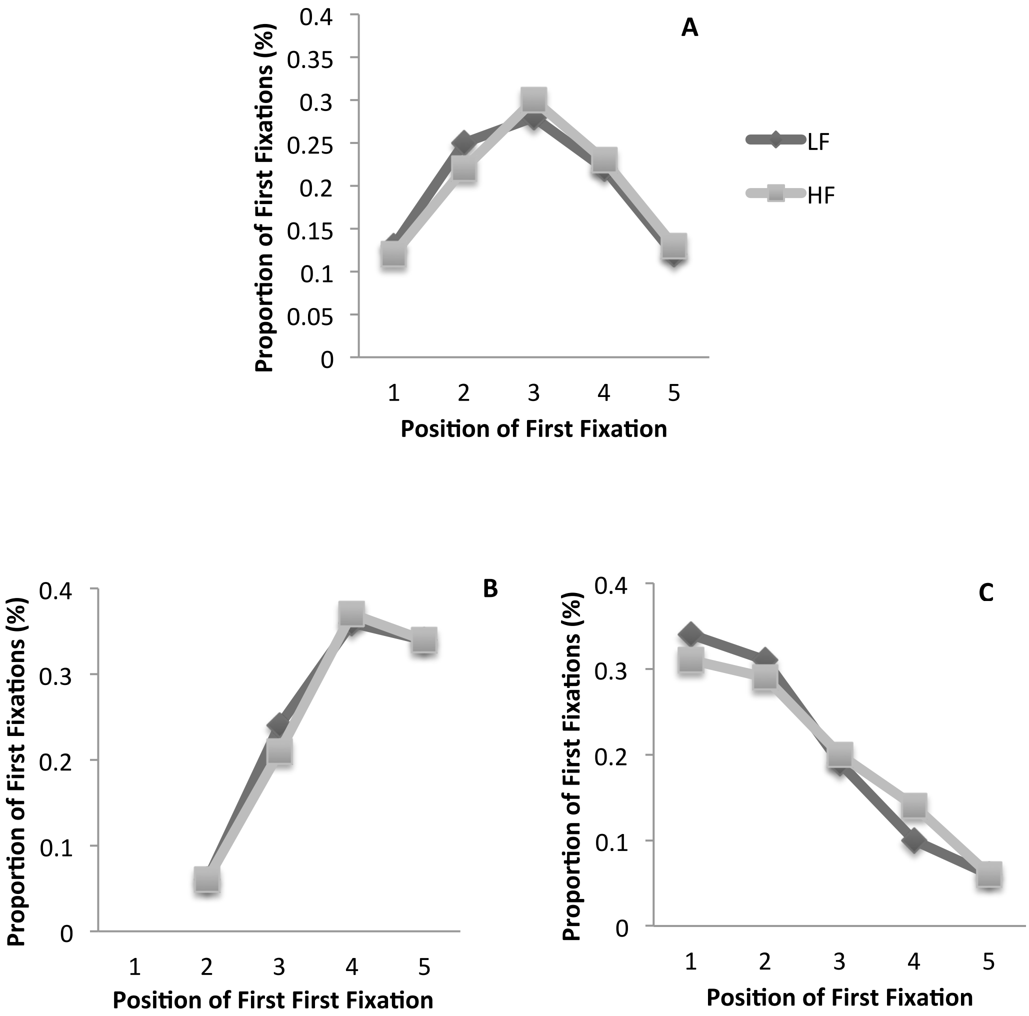


Figure 6.2. *Proportion of Fixations (%) aimed at the 1st to 5th Quintiles of the Noun when the Sentence either (A) Remained Static (B) Shifted to the Left, or (C) shifted to the Right.*

Left Shift: Unsurprisingly, a leftward shift resulted in later first landing positions on the noun (4.24 vs 2.93 character spaces: b=1.29; SE=0.08; t=17.04) together with an increased probability of it being skipped (37% vs. 17%: z=9.73; p<.001). Since saccadic targets will have been determined prior to the boundary being crossed, these effects can be interpreted as a consequence of (rather than a response to) the shift taking place. There was evidence, however, that the shift did disrupt the reading process, with refixation probability on the noun increasing (10% vs. 17%: z=2.39; p<.05) and inspection times becoming inflated after a shift had taken place. Although this increase was not significant in last fixation duration (228ms vs 231ms: b=3.87, SE=3.27, t=1.18), it was marginally significant in first fixation duration (230ms vs 235ms: b=5.08, SE=3.05, t=1.67) and significant across all other measures (single fixation duration: 234ms vs. 239ms: b=9.74, SE=3.30, t=2.95; gaze duration: 250ms vs 269ms: b=19.39, SE=5.60, t=3.46 and go-past time: 278ms vs. 296ms: b=20.17, SE=6.88, t=2.93). It is therefore clear that the simulated overshoot interfered with the typical reading process, slowing the onward progression of the eyes through text.

Table 6.2. *Fixation Time Measures (ms) and Skipping Probabilities (%) and First Landing Positions (character Spaces) for the Noun.*

	No Shift		Left Shift		Right Shift	
	LF	HF	LF	HF	LF	HF
First Fix	238	222	241	228	211	209
Last Fix	235	222	239	224	208	204
Single Fix	243	226	247	2330	217	213
Gaze	262	236	284	253	253	238
Go-Past	294	259	318	274	288	252
Skip Prob	14	19	35	38	12	16
Land	2.89	2.98	4.23	4.25	1.81	1.96
Refix Prob	13	9	20	13	22	17

While first landing position on the noun was unaffected by its frequency class (*b*=0.06; *SE*=0.06; *t*=1.06), high frequency nouns were associated with an increased probability of being skipped (25% vs 29%: *z*=2.56; *p*<.05) and if fixated, a reduced refixation probability (16% vs 10%: *z*=-4.16; p<.001) together with shorter inspection times (first fixation duration: 239ms vs 225ms: *b*=-15.18, *SE*=3.64, *t*=-4.17; single fixation duration: 244ms vs 229ms: *b*=-18.40, *SE*=3.74, *t*=-4.92; last fixation duration: 236ms vs 223ms: *b*=-14.34, *SE*=3.68, *t*=-3.90; gaze duration: 272ms vs 244ms: *b*-30.12, *SE*=5.50, *t*=-5.45 and go-past time: 305ms vs 266ms: *b*=-40.05, *SE*=9.30, *t*=-4.31). Importantly, there was no indication that whether or not a shift had occurred interacted with noun frequency in any measure (skipping probability: *z*=-0.22; *p*=.29; all other *t*s<1). Taken together, these results appear to suggest that if the reader overshoots their intended target, a stay and process strategy is adopted, allowing effects of noun frequency to be expressed in a similar fashion irrespective of shift condition.

Right Shift: Earlier first landing positions (1.88 vs. 2.93; b=-1.04, SE=0.09, t=-11.81) and a reduced probability of skipping (17% vs. 14%; z=-2.05, p<.05) were observed following a rightward shift. These two measures were also influenced by noun frequency, with high frequency words being skipped more often and receiving later first landing positions than low frequency words, although this latter effect was only marginally significant (13% vs. 18%; z=3.75, p<.001; 2.84 vs. 2.95 character spaces; b=0.13, SE=0.07, t=1.80). There was no evidence that shift interacted with frequency in either measure, which is unsurprising since any initial post-boundary "where" decision will have been determined prior to the shift taking place.

The earliest response to the sentence shift can be seen in first fixation duration, and here we see that the response to the rightward shift was markedly differently from the response following a leftward shift. First fixation durations were significantly shorter following a rightward shift compared with when the sentence had remained static (210ms vs. 230ms; b=-18.60, SE=3.61, t=-5.15) and while there was a main effect of frequency (b=-9.91, SE=2.90, t=-3.42), this was confined to the no-shift condition (high-222ms vs. low-238ms:

b=-16.36, SE=4.29, t=-3.82) with virtually no effect following a rightward shift (high-209 vs. low-211ms: b=-2.86, SE=3.21, t=-.89). This modulation of the frequency effect was reflected in a significant interaction (b=13.89, SE=4.90, t=2.84) and this can be seen in Figure 6.3. A similar pattern of effects was apparent in single fixation duration (shift: b=-18.56, SE=3.72, t=-4.99; frequency: b=-12.76, 3.23, t=-3.95; interaction: b=12.61, SE=5.08, t=2.48) and last fixation duration (shift: b=-21.46, SE=3.60, t=-5.96; frequency: b=-9.04, SE=3.01, t=-3.00; interaction: b=-8.84, SE=4.98, t=1.77).

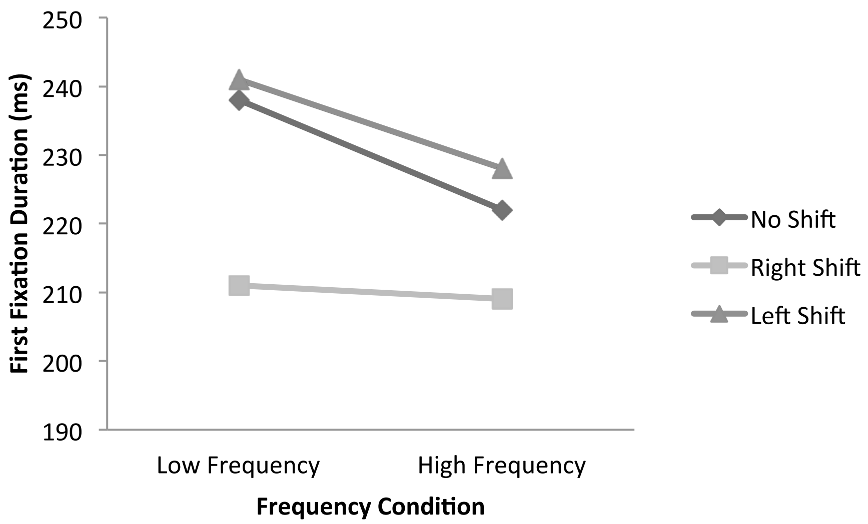


Figure 6.3. *Mean Duration of First Fixations (ms) Falling Within the Noun for Each of the Shift Conditions and as a Function of Noun Frequency.*

To investigate the potential mechanisms responsible for the differing response strategies following a left-shift and a right-shift compared to when the sentence remained static, first fixation duration was plotted as a function of initial landing position on the noun together with the probability of the noun being refixated. This is presented below in Figure 6.4 for each of the shift conditions.

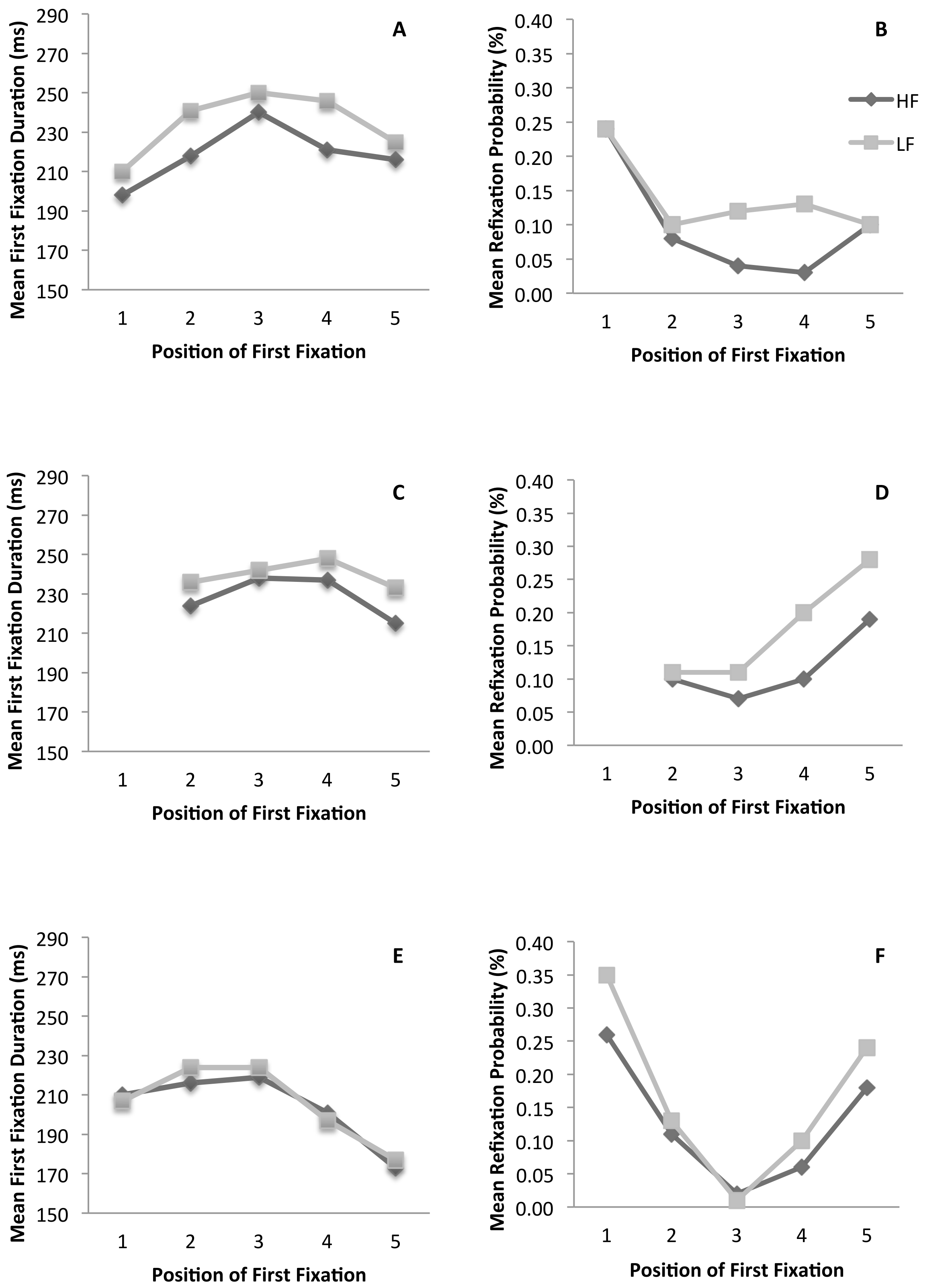


Figure 6.4. *Mean First Fixation Duration (left column) and Refixation probability (right column) for First Fixations Landing within Each of the Five Quantiles of the Noun as a Function of Frequency for the No Shift (A & B), Left Shift (C & D) and Right Shift (E & F) Conditions.*

First, it is worth noting that when the sentence remained static, the first fixation IVOP distribution as a function of frequency (Figure 6.4.A) bears a striking resemblance to that reported by Vitu et al (2001), with an apparent effect of frequency within each quintile. Vitu et al's results were based on a large corpus of eye movements, so this finding verifies the conclusion that this pattern reflects a phenomenon apparent in both naturalistic and 'experimental' reading studies.

The refixation probability distributions when the sentence remained static do, however, deviate somewhat from the pattern observed by McConkie et al (1989). McConkie et al reported similar refixation curves for low and high frequency words, with only the vertical offset indicating an increased tendency to refixate low frequency words. The pattern present in Figure 6.4.B, however, appears to show two different functions for high and low frequency words, with the high frequency words producing the typical U-shape function (e.g., McConkie et al, 1989; Vitu et al, 2001; Nuthmann et al, 2005), but with less of a dip near the centre for low frequency words.

This divergence in pattern might be related to differences between the two studies: While McConkie et al examined a large corpus of eye movements with a variety of word types and lengths (up to 8-letters) included in their analyses, the present set of results were limited to 5- or 6-letter nouns embedded within a selection of sentences that always adhered to a specific sentence structure. It is entirely possible that differences such as these might influence refixation rates, potentially producing differing floor effects for refixations. It could also be the case that McConkie et al found differing patterns for high and low frequency words, but since they did not report separate functions for high and low frequency words for words of differing lengths, this cannot be determined. What is apparent from the present set of results, however, is that, for 5- and 6-letter words, there appears to be a minimum natural refixation rate for low frequency words.

As discussed in the Introduction, it has been suggested that the IOVP may be related to the probability of initiating a quick error-correcting refixation, and this relationship has been used to explain why these two distributions typically mirror one another in an inverted fashion (Nuthmann et al, 2005; 2007). When the sentence remained static, the IOVP and refixation probability distributions for the high frequency words presented in Figure 6.4. appear to lend support to this proposal. To a lesser extent, so do the distributions for low frequency words, although as just discussed, these words appear to be associated with an overall higher probability of refixation, presumably triggered by increased processing difficulty associated with the low frequency nouns employed in this study. A similar relationship between the IOVP and refixation distributions is also apparent following both the left and right sentence shifts. These distributions therefore lend support to the suggestion that the IOVP may be driven by quick error-correcting saccades.

The distributions of refixation probabilities (Figure 6.4, right panel) for the two shift conditions suggest that first landing position is a strong predictor of whether a word will be refixated, with this function – showing minimum refixation probability at a word's centre - surviving both left and right sentence shifts. These distributions therefore clearly indicate that refixation probability is a reflexive process based on within-word landing position. This is also consistent with the finding that refixation probability distributions survive when participants engage in 'mindless reading'. That is, when readers are instructed to read a row of x's punctuated by spaces that imitate sentences but which are devoid of lexical information. Thus this pattern appears to indicate an oculomotor, rather than lexical, basis for these refixation curves (Vitu, O'Regan, Inhoff & Topolski, 1995; Nuthman et al 2007). Taken together, these results provide convincing evidence that whether one refixates a word is not necessarily linked to lexical processing; rather first landing position within a word also appears to make a substantial contribution.

The IOVP function for the left-shift (overshoot) condition is similar in form to the no-shift condition, with evidence for a frequency effect at every position, but with the peak shifted rightwards, suggesting the response to the simulated overshoot was to stay and process the noun from the unintended location. Following a right shift, however, there is an apparent flattening of the IOVP function, a lack of a frequency effect at all positions, which combined with the lower overall first fixation durations on the noun, suggests that the simulated undershoot tended to trigger a quick error-correcting saccade.

It could, however, be argued that the apparent frequency effect in refixation probability for fixations landing on the peripheral regions of the noun following a simulated undershoot does not necessarily support the quick error-correction hypothesis; indeed, they imply a degree of lexical processing had been engaged. This is also apparent in a clear effect of frequency in refixation probability for the overall noun region, with more refixations when the noun was of a low compared to a high frequency: (18% vs. 13%: $z=-3.76$, $p<.001$), and while there was a main effect of shift, with more refixations following a rightward shift (19% vs. 11%: $z=4.16$, $p<0.001$), there was no evidence of an interaction between these two variables ($z=0.65$, $p=.52$).

However, this finding is not necessarily incompatible with the quick error-correction account of saccadic undershoots. The frequency effect in quintile 1 could arise if a quick error correcting saccade had been programmed, but subsequently cancelled on those occasions that lexical processing was sufficiently advanced before the refixation was committed to action. According to McConkie et al (1988), a first landing position on the OVP indicates a saccade has successfully landed on its target. In the present experiment, average saccade length into the noun was 8.5 characters. It is not surprising that this is slightly longer than the average of 7-character spaces reported by McConkie et al since the present experiment was designed to trigger longer than average fixations into the noun, achieved by keeping the verb relatively long with redundant endings. Given that information can be extracted from words up to 14-character spaces to the right of fixation (McConkie & Rayner, 1976), it is likely that a fixation 8.5 character spaces away will have permitted some preview of the noun. Thus even if the optimal viewing position was not fixated, due to the simulated undershoot, the lexical processing of high frequency nouns might have been sufficiently advanced before the labile stage of saccadic programming expired, allowing the refixation to be cancelled and a new saccadic target to be identified. In contrast, a low frequency word fixated in a non-optimal location may not, due to a combination of acuity constraints and lexical difficulty, have had its processing advance at a fast enough rate to cancel the refixation before it was committed to action.

A similar scenario could also apply to the presence of the frequency effect within quintiles four and five. These fixations will have been aimed at the first characters of the spillover region, but fallen short due to the simulated undershoot. This would be the case if the intention had been to skip the noun. The most likely reason for that would have been if lexical processing on that noun was relatively far advanced. Thus, an automatic refixation might have been programmed upon fixating a non-optimal location on the noun, but that fixation could be cancelled if the word was sufficiently processed before the non-labile stage committed the refixation to action; a scenario far more likely If the noun was high frequency.

The lack of a frequency effect in quintiles two and three is not surprising since, in these cases, the saccade will have been aimed for the end of the word, implying lexical processing on the noun was already well advanced. Thus by the time the OVP was (accidently) fixated, any confirmatory processing during the fixation could be completed exceedingly quickly, resulting in the absence of a frequency effect. Overall, therefore, the refixation distributions here following a rightward shift are not incompatible with the hypothesis that saccadic undershoots normally tend to be followed by the programming of a quick error-correcting refixation.

Perhaps the cleanest way to investigate the response to undershooting a word is to compare fixation durations on the last two character spaces of the verb preceding the noun when (a) a fixation had been intended to land on this region (i.e. there was no shift) and (b) the fixation had been intended to land on the noun, but due to (simulated) saccadic error (i.e., following the rightward shift), it had fallen short. If a stay-and-process response occurs with an undershoot, we should observe a frequency effect from the noun on the duration of fixations following a simulated undershoot. If however, a quick error-correcting saccade is initiated, as the results above seem to suggest, then we should observe no frequency effect following a simulated undershoot within this region.

To test this, a new two-character region – the last two character spaces of the verb - was defined. First fixations durations falling within this region were compared depending on whether they had been intended to land here (i.e., the sentence had remained static), or whether they had been intended to fixate the beginning of the noun (i.e., they landed on the final two characters of the verb following a rightward text shift).

As can be seen from Figure 6.5, these data provide convincing evidence against the claim that an accidental undershoot of a word is likely to be associated with a stay and process strategy (Rayner et al, 2004; Drieghe et al, 2008) rather than a quick error-correcting saccade (Nuthmann et al, 2005; 2007). As the Figure illustrates, there was a numerical trend for fixations following a rightward shift to be 11ms shorter than if the sentence had remained static (b=-11.79, SE=9.59, t=-1.23). Furthermore, there was no evidence that shift interacted with frequency (b=11.63, SE=14.97, t=0.78). Thus, Irrespective of whether the saccade had intentionally landed on the end of the verb or not, it does not appear that the noun was being lexically processed while the eyes were directed towards the end of the verb (frequency: b=5.56, SE=6.49, t=-0.86). Indeed, the 4ms numerical trend in the frequency effect following a simulated undershoot fell in the opposite direction to what would be predicted from a stay and process response.

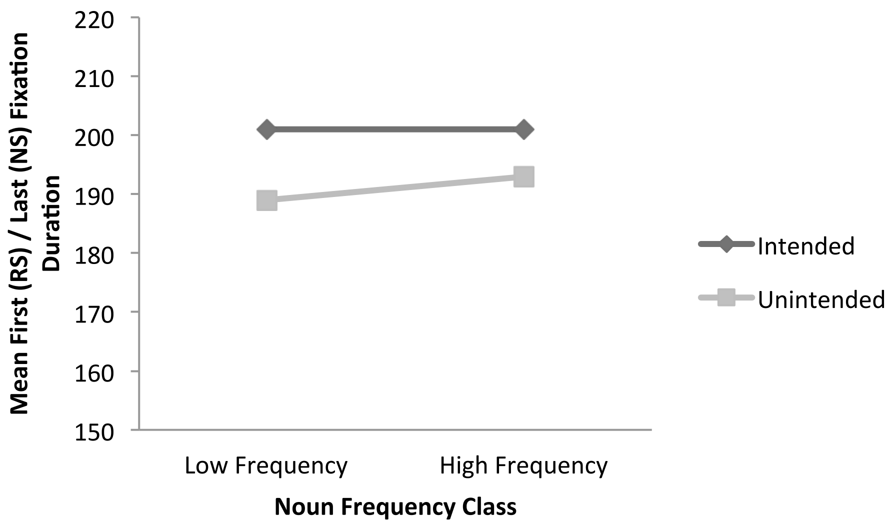


Figure 6.5. *Mean First Fixation Duration (Right Shift) and Last Fixation Duration (No Shift) for Fixations Falling on the Last Two-Characters of the Noun.*

Despite the simulated undershoot (right shift) causing shorter first fixation durations on the noun and eliminating the frequency effect on it, later measures show that this initial disruption was only temporary. There is an absence of an effect of shift in both gaze duration and go-past time (b=-2.93, SE=3.63, t=-0.81 and b=-6.56, SE=5.64, t=-1.16, respectively), and a clear effect of frequency in both (low-257ms vs. high-237ms: b=-21.39, SE=4.06, t=-5.27 and low-291ms vs. high-256ms: b=-36.29, SE=7.84, t=-4.63, respectively). The marginally significant interaction between shift and frequency in gaze duration (b=12.79, SE=6.62, t=1.93) seemed to reflect a marginally stronger effect of frequency in the no shift condition (b=-26.08, SE=6.24, t=-4.18) compared to the right shift condition (b=-14.33, SE=5.10, t=-2.81), with any evidence for an interaction disappearing completely when the later measure of go-past is considered (b=-.07, SE=11.15, t=-0.06).

6.3.3. Spillover Region

Left Shift: Earlier first landing positions within the spillover region were observed following a leftward shift (4.49 vs. 4.96 character spaces; b=0.47, SE 0.19, t=2.46). There was, however, no evidence of a frequency effect in first landing position (b=-0.01, SE=0.12, t=-0.06) or a shift by frequency interaction (b=-0.07, SE=0.22, t=-0.33). If the effect of shift had been a consequence of the sentence shifting (i.e., this was the first post-boundary fixation) rather than a response to it, then an opposite pattern of effects would be predicted mirroring those seen on the noun (see above). The direction of the effect, combined with the finding that there was no evidence for lexical processing difficulties of the noun spilling over into the spillover region for this measure indicates that the earlier first landing positions following the left shift might simply reflect a more cautious reading style being adopted after a shift had taken place.

Table 6.3. *Fixation Time Measures (ms) and First Landing Positions (character Spaces) for the Spillover Region.*

	No Shift		**Left Shift**		**Right Shift**	
	LF	**HF**	**LF**	**HF**	**LF**	**HF**
First Fix	213	211	209	208	206	203
Gaze	374	382	370	380	383	368
Go-Past	423	419	478	456	415	401
Landing	4.48	4.45	3.97	4.05	4.51	4.50

There was no evidence for a frequency driven spillover effect in first fixation duration (b=-1.16, SE=2.71, t=-0.43), gaze duration (b=-9.06, SE=6.97, t=1.30) or go-past time (b=13.70, SE=10.44, t=1.31). There was a delayed cost of the sentence shifting, evident in go-past time (no shift: 421ms vs. left shift: 467ms; b=45.92, SE=13.28, t=3.46), but no evidence that the sentence shift affected either first fixation duration (b=2.84, SE=3.03, t=0.94) or gaze duration (b=-2.12, SE=8.48, t=-0.25). There was no evidence that shift and frequency interacted in any of the reported measures (all ts<1). It appears, therefore, that while the simulated overshoot encouraged a more cautious reading style in terms of where fixations were located, apart from an occasional tendency to regress, it did not appear to affect the onward progression of the eyes. Apart from increased caution, the consequences of the overshoot appear to be primarily reflected in fixations falling on the noun, which is consistent with the proposal that any significant difficulties caused by the shift were tackled and resolved within that region.

Right Shift First fixation durations were significantly shorter following a rightward shift (205ms vs. 212ms; b=7.29, SE=2.58, t=2.83). Since this effect was still present following a fixation of the noun (206ms vs 214ms: b=7.58, SE=2.76, t=2.75), it cannot be the case that it was simply the product of it being the first fixation after the sentence had shifted (i.e., it was the first fixation after the noun had been skipped triggering a similar response to that seen on the noun). Instead, it appears that an undershoot does influence the 'rhythm' of the eyes, producing shorter fixations both on the noun and when first entering the spillover region. This disruption is however, only temporary, since the rightward shift failed to significantly influence any other measure (gaze duration: b=-4.26, SE=6.76, t=-0.63; go-past time: b=-14.43, SE=9.02, t=-1.60; first landing position: b=-0.07, SE=0.14, t=-.50). Similarly to a leftward shift, there was no suggestion of a frequency-driven spillover effect in first fixation duration: b=-2.72, SE=2.96, t=-0.92; gaze duration: b=2.22, SE=6.13, t=0.36; go-past time: b=9.95, SE=8.52, t=1.17; or first landing position: b=0.03, SE=0.11, t=0.27), nor was there any evidence that shift interacted with frequency (gaze duration: b=-20.80, SE=14.91, t=-1.40; go-past time: b=9.73, SE=17.01, t=0.57; first landing position: b=0.00, SE=0.21, t=0.02). Thus, aside from immediately shorter first fixation durations, there was no evidence for any continuing adjustments to reading strategy following a simulated undershoot. Again, the disruption caused by the sentence shift appears to have been primarily localised to the noun.

6.4. General Discussion

The concept of a mislocated fixation has been invoked in two very different ways to explain two very different phenomena. While Nuthmann et al (2005; 2007) hypothesise that mislocated fixations are followed by a quick error-correcting saccade and therefore responsible for the IOVP effect, Rayner et al (2004; Drieghe et al, 2008) suggest that mislocated fixations followed by a stay and process strategy are responsible for the expression of lexical parafoveal-on-foveal effects. But as Kliegl and Engbert (2011) point out, these two positions are mutually exclusive. To help differentiate which of these two possibilities reflects the true oculomotor response, the present study sought to investigate how the oculomotor system responds to mislocated fixations using the text shift paradigm.

Previous research has consistently shown that word frequency has a very early impact on the eye movement record, with effects of frequency routinely observed on the duration of the first fixation on a word (e.g. Inhoff, 1984; Schilling et al, 1998; Rayner, Ashby, Pollatsek & Reichle, 2004; Angele & Rayner, 2011). This made frequency the ideal lexical property to manipulate in order to determine the time course of lexical processing on the noun immediately after a shift had taken place. When the sentence remained static, there was clear evidence for the well-replicated frequency effect, which was consistently expressed across all durational measures. And while there was an increase in all durational measures following a left shift – highlighting a temporal cost to overshooting - there was no evidence that shift modulated the frequency effect, which was again clearly present across all durational measures on the noun. This indicates that overshooting one's target does not delay the onset of lexical processing on the noun, suggesting a stay and process strategy had been adopted in this case. It is clear, however, that this strategy is associated with a cost, with increased durations following the simulated overshoots; potential causes for this will be returned to below.

The immediate response to a simulated undershoot was markedly different. First fixations on the noun following a simulated undershoot were shorter than when the sentence had remained static and there was a complete absence of a frequency effect. These results appear to suggest that a quick error-correcting response to the mislocated fixation had been engaged. The flattening of the IOVP function following a rightward shift combined with the lack of frequency effect at every quintile supports this hypothesis. Thus, it seems that the response to an undershoot is to delay lexical processing until the eyes relocate to a more optimal location. Since the effect of frequency was present in all cumulative measures on the noun, and since these measures failed to reveal any prolonged effects of shift, this indicates that the initiation of a quick error correcting saccade was an efficient strategy for maintaining the mean rate of eye movements through text and the continuation of effective lexical processing.

It is important to reiterate that these shifts only perturbed the eye by two character spaces. Thus the majority of fixations contributing to the reported measures will not have been ‘mislocated’, in the sense that the saccade had missed its intended target word. Rather, there was simply a degree of error regarding where the eye landed within the targeted word. It was somewhat surprising therefore, to observe such a specific set of responses to such minor errors, with the chosen response apparently driven by whether the saccade under- or overshot its intended target location, rather than being driven by whether it landed in the correct word.

This is not the first time that an asymmetrical response to mislocated fixations has been reported in the literature following minor simulated saccadic errors. Feng (2009, Experiment 3), who shifted text left or right by one to three character spaces, reported more regressive fixations following a simulated overshoot but no such differences following a simulated undershoot. While the present study did not analyse the proportions of different saccade types following a sentence shift, the specificity of the response, with increased go-past times following a simulated over- but not undershoot would also seem to reflect a hesitance in moving forward immediately after a saccade has over-, but not undershot its target.

The present sets of results also converge with those of Feng (2009) with respect to fixations that undershot their intended target word, landing instead at the end of the preceding word. Feng reported that, when the subsequent saccade was either progressive or a refixation (which together accounted for approximately 90% all post-shift saccades), first fixation durations were shorter following a rightward shift than when the sentence had remained static; a finding that he interpreted as reflecting a potential error correction mechanism. The present study did not separate the data based on saccade type, but it did show a similar numerical trend with shorter first fixation durations following a saccadic undershoot that positioned the fixation at the end of the verb. While the 10ms effect was non-significant in the present study - presumably due, in part at least, to the reduced power in the analysis[41] – it is interesting that the numerical pattern follows that observed by Feng.

Importantly, when a mislocation caused the reader to undershoot the noun and land instead on the end of the verb, there was no evidence, statistical or otherwise, to suggest that the parafoveal word had been lexically processed during that mislocated fixation, with clear absence of any effect of the noun's frequency within this verb-end region. What is clear from this set of results, therefore, is that it seems extremely unlikely that saccadic undershoots are responsible for the expression of lexical parafoveal-on-foveal

[41] The number of fixations contributing to this analysis were attenuated since (a) the region was only two characters in length and (b) as Figure 6.5 indicates, this position is associated with fewer fixations than mid word regions. Fewer fixations therefore led to reduced power in the analysis.

effects, as suggested by Rayner et al (2004) and further advocated by Drieghe et al (2008). Whether or not this necessarily implicates the sort of error correction strategy proposed by Nuthmann et al (2005; 2007) will be returned to below.

The results of this study converge with and extend those of Feng (2009) in suggesting that the oculomotor response to saccadic error is dependent upon whether a saccade under- or over-shoots its intended target. Taking a moment to reflect, it makes sense that different response strategies should be engaged: While an undershoot temporarily confines a saccade to a location in which information had presumably already been extracted, an overshoot requires the reader to ‘fill in the gap’ caused by the overshoot, which could be addressed either cognitively (e.g., by guessing) or perceptually (by means of orthographic extraction from the missed region). It therefore makes sense that an undershoot is followed by a fast-acting corrective saccadic to allow the eyes to continue moving forward at the preferred rate in order to achieve efficient lexical processing, while an overshoot will require a slowing of the system to allow perceptual and/or cognitive processing to catch-up.

Assuming that such gaps are ‘filled in’ with slower extraction of orthographic information allowing for the expression of main effects of shift and frequency, this also explains the lack of interaction between the two factors, since altering rates of orthographic extraction should precede and therefore not interact with, the time required to lexically process a word (e.g., Reingold & Rayner, 2006). If the missed information had been 'filled in' cognitively, one might expect an interaction with higher frequency words being easier to guess. Thus, the former explanation seems the most plausible, given the present set of results.

That is not to say that error correcting saccades are confined to saccadic undershoots; indeed, both the present study and Feng observed increased refixations and regressions, respectively, following a simulated overshoot. However, it is possible that such a strategy is only invoked after a degree of effort has been made to stay and process from the erroneous location. Planning and executing a regression due to oculomotor error takes time: first, assuming the error is detected via retinal input (as these data and those of Feng suggest), a decision to relocate must be delayed by a time proportional to the eye-mind lag, then the saccade must be programmed and executed. This could potentially make it a less efficient option than staying and processing from a suboptimal location. Indeed, this may especially be the case if the text is relatively 'easy' to process (i.e., the words are highly predictable, frequent or short), and the 'gaps' in the text can be either cognitively or perceptually 'filled in' with relative ease from a sub-optimal location. The initiation of a quick error-correcting saccade in these cases might result in the gap being filled after the regression has been committed to action, making the error-correction strategy both a redundant and time costly exercise.

Overall, it seems that for both saccadic under- and overshoots, the oculomotor system acts to prioritise efficiency, and this induces different responses depending on processing requirements: the first is to immediately engage a quick error correcting response, while the latter – due to the trade-off between word processing time and the time required to relocate – might only engage the error correcting strategy once processing from the erroneous location looks as though it might fail.

The strategy employed here on the noun to cope with a simulated undershoot appears to have been relatively successful in resolving any disruption that the sentence shift had caused. The only delayed effect in the spillover region following an undershoot was shorter first fixations, suggesting that the simulated undershoot temporarily encouraged a quicker succession of fixations, presumably to prevent the eyes from falling behind lexical processing. This strategy was clearly temporary since there was no evidence that shift modulated any of the cumulative measures in the spillover region.

Following an overshoot, however, the earlier first landing position appears to reflect a more cautious reading strategy, keeping the eye close to the region of disruption longer than it otherwise would have. It also appears that there was an increased tendency to regress from the spillover region following a simulated overshoot, evidenced by an absence of an effect of shift in gaze duration but a clear effect in go-past time, which includes regressions. This appears to suggest that the stay and process response was not always successful, perhaps because the reader had misperceived or misidentified the previous lexical entry due to attempting to stay and process from a sub-optimal location, but this error in processing might only have been realised after subsequent text had been read. But what is clear from the lack of a frequency spillover effect or frequency x shift interaction in the spillover region is that the lexical processing of the noun was confined to the noun and inducing mislocated fixations did not alter this.

One result that does appear to contradict the assumption that lexical processing of the noun was confined to the noun is the finding that refixation probability was 6% higher for low compared to high frequency words on quintile 5 of the noun following a simulated undershoot. This is a region where the programmed saccade appeared to have been intended to fixate the first word in the spillover region. This seems to imply that such saccades had been destined for the spillover region prior to the full lexical processing of the noun. While not necessarily at odds with a parallel model of eye movement control, this finding is particularly difficult to reconcile with a model – such as the E-Z Reader model - that assumes successful lexical identification is necessarily the engine driving the eyes through text.

There are, however, two means by which E-Z Reader might account for these results. First, the trigger to plan a saccade to the spillover region should, according to the E-Z Reader framework – have been completion of the L1 stage of lexical processing on the noun. It might therefore have been the case that occasionally, stage L2 had not been completed on the low frequency noun by the time the saccade was executed, permitting an effect of frequency to be expressed on the saccadic undershoot condition in quintile 5. There is, however, one result that seems at odds with such an interpretation: this above explanation hinges on the same mechanism used to explain frequency driven spillover effects, and there was no evidence for a frequency-driven spillover effect in first fixation duration in the spillover region. The only other explanation for the frequency-driven refixation effect in quintile 5 following a saccadic undershoot, that is compatible with the E-Z Reader model, is that these fixations represent cases of a double-saccadic error; that is, they had erroneously been targeted at the beginning of the spillover region, but the induced saccadic error cancelled this error out and resulted in the saccade landing precisely where it was meant to. This explanation is potentially feasible but in the present study untestable and therefore difficult to evaluate. This explanation does seem somewhat tenuous, however, with the most parsimonious explanation being that a looser coupling between fixation location and attention exists; or more precisely, that a looser relationship between fixation location and lexical identification exists.

In addition to not spilling over into the spillover region, there was also no evidence that noun frequency influenced fixation durations on the preceding verb. Initially, this result was somewhat surprising given that, according to previous research, the right conditions were present for frequency driven parafoveal-on-foveal effects to be expressed; namely, the foveal word was relatively long and the parafoveal word was relatively short (Hyönä and Bertram, 2004). Closer inspection of the materials does, however, provide an explanation for the lack effect. Namely, the materials were designed to encourage a fixation towards the beginning of the noun, and relatively few fixations towards the end of the verb. This was achieved by employing high frequency verbs with redundant endings (e.g., “ing” and “ed”). This manipulation appeared to be successful with an average of just 1.12 fixations on the 7- or 8-character verb, which landed, on average, on letter position 3.12. It is likely therefore that too few fixations fell towards the end of the verb to allow for the expression of a lexical parafoveal-on-foveal effect – and this will tend to inhibit potential ‘parafoveal effects’, irrespective of whether one assumes serial or parallel lexical processing.

Finally, the present set of results clearly provides strong evidence regarding how mislocated fixations are detected. If the initial response to a mislocated fixation had been based on an efference copy, then there should have been no difference between the no shift, left or right shift IOVP distributions, since the shift will have occurred after the point in time in which the efference copy will have been written. The clear differences between these conditions therefore suggests that mislocated fixations are detected on the basis of retinal input and not an efference copy. Feng reached the same conclusion after finding a clear IOVP distribution on a word when first fixation durations on that word were plotted as a function of their first landing post-shift position, but not their first landing pre-shift position. The present set of results thus aligns with Feng (2009; Experiment 3) to suggest that mislocated fixations are detected on the basis of retinal input rather than an efference copy.

As already discussed, the pattern of results obtained in this study appear to violate what has become an important tenet of the E-Z Reader model: that saccadic undershoots will frequently be followed by a stay and process response. Indeed, there was no evidence for such a response to saccadic undershoots irrespective of whether the saccade missed the intended target word or whether the mislocation simply caused an error in where within the word the eye landed. This finding alone carries important implications for the E-Z Reader model, since the stay and process response to mislocated fixations has, to date, been the only explanation proponents of this model have been able to propose for the apparent expression of lexical parafoveal-on-foveal effects. Despite clearly failing to correctly predict the response to saccadic undershoots, such a stay and process response does appear to operate following saccadic overshoots, but this is not currently implemented in the model.

As will be recalled from Chapter 1, the E-Z Reader model (Version 6 and onwards) is capable of simulating mislocated fixations and correctly assumes that they are detected based upon retinal input. In the current instantiation of the model, within-word mislocations act to modify (a) lexical identification times and (b) refixation rates, both of which increase with increasing eccentricities. Since the present research indicates a specific set of systematic asymmetric responses to mislocations that are currently not specified within the model, it will be interesting to see whether it is possible for the model to incorporate mechanisms reflecting these responses without jeopardising its current capabilities in accounting for the wealth of benchmark findings. Of course, it will be of particular interest to see whether the model can adapt to the finding that saccadic undershoots are not typically followed by a stay and process strategy and whether an alternative explanation for lexical parafoveal-on-foveal effects can be found.

The SWIFT model, in its current instantiation, also fails to capture the dynamic response to mislocated fixations. First, the model assumes that mislocated fixations are detected based on an efference copy, and as just discussed, the evidence here, as well as that provided by Feng (2009) suggests that is not the case. Like the E-Z Reader model, SWIFT assumes that saccades target word centres and that processing rates are adjusted for saccades that miss this target, with increased processing rates associated with increasing eccentricity. The model does not, however, incorporate any oculomotor responses to deal with within-word errors. The SWIFT model does incorporate a mechanism that deals with mislocated fixations that miss their intended target words, although the model predicts a quick error correcting strategy following both saccadic under- and overshoots, which appears to be in conflict with the present set of results. To accommodate these results, future versions of the SWIFT model would need to (a) allow for an asymmetrical response to within word saccadic error and (b) delay the response to a mislocated fixation to allow for the eye-brain lag to determine that the fixation was indeed mislocated.

Thus, while the E-Z Reader model includes a mechanism to respond to intra-word errors (via refixations) and the SWIFT model incorporates a mechanism to deal with inter-word errors (via fast error-correcting saccades), neither model captures the dynamics or the nature of the systematic, asymmetric, response observed here.

Finally, the patterns of results presented here do not appear to be consistent with Nuthmann et al's (2005; 2007) explanation for the IOVP effect. According to Nuthmann et al, a quick error-correcting response is initiated when a saccade lands on an erroneous word. This strategy, they suggest, results in shorter first fixation durations at the peripheral regions of a word, compared with the central regions where mislocated fixations of this nature will be less prevalent. The results of this experiment, however, have shown that the quick error-correction strategy is initiated for within word, as well as between word errors. Furthermore, the quick error-correcting response appears to be most clearly apparent with fixations that under- rather than overshoot their intended target. As discussed above, the primary motivation for this experiment was to ascertain if mislocated fixations could be responsible for the expression of parafoveal-on-foveal effects and semantic preview effects; it was not designed to test Nuthmann et al's account of the IOVP effect. It appears that Nuthmann et al were correct to attribute some aspects of the IOVP to quick error correcting saccades, but a full account of the IOVP effect needs to take into account the asymmetry of response and the fact that corrections also appear to be triggered by erroneous within word landings.

6.5. Conclusion

Using the text shift paradigm, it was possible to investigate the time course of lexical processing on a critical word immediately after a saccadic under- or over-shoot had taken place. The results implicate an asymmetric response to these errors, apparently driven by the need to engage the most efficient strategy to resolve any disruption and re-establish efficient lexical processing. This response of course differs depending on whether the saccade under- or overshot its intended target.

While efforts have been made to simulate mislocated fixations in both the SWIFT and E-Z Reader models, both appear to have greatly underestimated the complexity of the response. Given the estimated frequency of mislocated fixations during reading (Nuthmann et al, 2007) and their ability to prompt such systematic responses, it is imperative that future research is dedicated to further investigating how the oculomotor system responds to these errors and that the outcome of this research is reflected in future versions of the models; the text shift paradigm will undoubtedly become a valued tool in this endeavour.

But for the moment it seems clear that saccadic undershoots resulting in a fixation which falls on the end of the preceding word, or even the first few characters of the intended word, tend to trigger a quick error-correction and are not associated with a stay and process response.

CHAPTER 7

General Discussion

7.1. Introduction

This thesis sought to investigate the link between fixation location and attention during reading. To this end, three lines of enquiry were followed. First, the question of whether parallel lexical processing is psychologically plausible was addressed (Experiment 1). Second, the nature and extent of parafoveal processing during reading was investigated via a series of gaze contingent display change experiments (Experiments 2-6). And finally, mislocated fixations were experimentally induced in order to assess whether oculomotor error together with a stay and process response might account for seemingly parallel effects within a serial framework (Experiment 7). The main themes to have emerged from this research will be discussed in this chapter. First, however, a brief reflection on how computational models have influenced the current research climate will be considered.

7.2. Models of Eye Movement Control During Reading

The number of citations containing the words “serial”, “parallel”, ”eye movement” and “reading” has grown exponentially since the introduction of the E-Z Reader model in 1998. Prior to this, fewer than 10 citations per year contained these words. In contrast, by 2012 the number of citations had risen to 108 per year (Murray, Fischer & Tatler, 2015). These figures highlight just how instrumental the E-Z Reader model has been in generating research concerned with the distribution of attention during reading. The E-Z Reader model has also been the driving force in the creation and on-going development of opposing models of eye movement control, such as SWIFT in 2002 and Glenmore in 2006.

These models have been instrumental in altering the experimental approach researchers take in investigating the distribution of attention during reading. Through their ability to make quantitative predictions, they have encouraged researchers to focus on specific sets of phenomena that are considered to distinguish between them, such as semantic preview effects, parafoveal-on-foveal effects and word n+2 preview effects. It is unlikely that researchers would have persevered in drilling for such effects without the potential ramifications such effects might have for the continued viability of the models.

It is evident from even a brief scan of the literature that, over the past fifteen years, an immense amount of research has been centred on testing the predictions of the E-Z Reader model; far more so than any other model. This thesis is no exception. While this unbalance might be considered to expose a bias, it should not be viewed as such. Quite the contrary, it is a testament to the utility of the E-Z Reader model that is has generated such a wealth of research. The model's mechanisms are coherently defined, making it accessible to the non-modelling expert. It is its ability to make transparent predictions that make it the perfect model for experimental psychologists to test.

A word of caution has, however, recently been raised by Rayner (2009). He highlighted the importance of testing a model's predictions via simulations rather than speculating on predictions that might be based simply on their architectures. Indeed, in a recent series of simulations conducted by Schotter et al (2014), it was shown that the E-Z Reader model is capable of simulating small word n+2 preview effects and semantic preview effects. However, for the reasons outlined in Chapter 5 of this thesis, further simulations appear to be necessary in order to evaluate the E-Z Reader model's capabilities in accounting for such effects. Nevertheless, this research does drive home the importance of testing the models thoroughly via quantitative simulations, rather than basing predictions on mere assumptions concerning the sorts of effects that a particular type of architecture will deliver.

Box (1979, p202) famously stated: "…all models are wrong, some are useful". Indeed, it is clear that the models discussed in this thesis are in no way direct replicas of the human cognitive processing system; they do not profess to be so. There is, however, no denying the utility of the models in highlighting how some apparently serial effects can be explained within a parallel architecture, while other seemingly parallel effects can be consistent with serial processing. To this end, these models have proved themselves to be invaluable tools in driving research on the link between fixation location and attention during reading.

7.3. Parafoveal Preview Effects During reading

It has long been known that the uptake of information from the written page is not restricted to the word being fixated. Words in the periphery that precede the fixated word, and those that follow it, have both been shown to influence foveal inspection times (e.g., Rayner & Duffy, 1986 and Starr & Inhoff, 2004, respectively). Furthermore, the properties of the fixated word have also been shown to modulate the degree to which this parafoveal information is extracted (e.g., Henderson & Ferreira, 1990). The past forty years investigating the interactions between foveal and peripheral word processing has been bedevilled with inconsistent results; the results of this thesis being no exception. This section will consider what these patterns of effects can tell us about the link between fixation location and attention during reading.

There is little controversy concerning the existence of word n+1 orthographic preview benefit. These effects have been shown in previous research (e.g., Rayner, 1975), and have been found consistently throughout this thesis (Experiments 2, 3, 4 and 5). And despite early scepticism regarding the existence of orthographic parafoveal-on-foveal effects (e.g., Rayner, White, Kambe, Miller & Liversedge, 2003), it is now generally accepted that these effects do exist, and again, such effects have been reported here in Experiments 2 and 3; the mechanisms that are potentially responsible for these effects are discussed below. A slightly more contentious issue, however, is whether orthographic pre-processing can occur on a word two-words downstream. While there is some evidence that this may be the case, effect sizes are typically small and are only occasionally reliable (e.g., Angele & Rayner, 2011; Radach et al, 2007; 2013).

This thesis raised and tested the possibility that the word length and frequency of words n, n+1 and n+2 might be responsible for the inconsistent nature of word n+2 preview effects. For example, low frequency and/or long words might restrict word n+2 effects by either absorbing all available attentional resources, or by pushing word n+2 outwith the effective span of apprehension.

The results reported here were somewhat surprising. First, it was anticipated that, in accordance with the word-grouping hypothesis (Radach, 1996, cited in Drieghe, Pollatsek, Staub & Rayner, 2008), word n+2 pre-processing should be enhanced when the word preceding it was a determiner rather an alternative high frequency word. The results of Experiments 2 and 6 did not support this assumption. In fact, when word n+1 was restricted to 3-letters, the strongest evidence for word n+2 pre-processing arose when word n+1 was an alternative high frequency word, not a determiner, with a word n+2 preview effect observed on word n (Experiment 2). As previously discussed, Kliegl et al (2007) also report a sensitivity to word n+2 preview restricted to cases where word n+1 was an alternative high frequency word. The most plausible explanation for this pattern of effects is that readers are more likely to seek information from word n+2 when the detailed interpretation of word n+1 depends upon the meaning of word n+2. That is, of course, not to say word n+2 pre-processing cannot occur when word n+1 is a determiner. Word n+1 type did not modulate the word n+2 preview effects in skipping probability on words n and n+1 in Experiment 2. Also, Radach et al (2013) reported word n+2 preview effects when word n+1 was exclusively a determiner. It appears therefore, that while word n+2 preview effects do occur when word n+1 is a determiner, they are perhaps more prevalent when the intervening word is an alternative high frequency word that requires the identification of word n+2 before it can be fully interpreted.

Second, there was a broader range of word n+2 preview trends and effects in Experiment 3, where word n+1 length was an average of 5-letters, rather than the 3-letters in Experiment 2. Furthermore, there were stronger word n+2 orthographic preview effects on word n+2 in Experiment 5 (where word n+1 was 4-letters) compared with Experiment 6 (where word n+1 was just 3-letters). The finding of increased word n+2 pre-processing when word n+1 was longer than 3-letters fits with the explanation for why these effects are more prevent for alternative high frequency words compared to determiners. Readers might have learnt that 3-letter words are often devoid of semantic information given the high frequency of determiners in text. However, the longer the word, the more likely it is that that word will carry semantic information, which, in the case of adjectives and other modifiers, will often need to be linked with semantic information derived from the immediately adjacent words.

These results, coupled with this explanation, raise the question of whether individual word units are necessarily the most appropriate unit of analysis if we wish to understand the way the eyes progress through text. Given the interactions between foveal and parafoveal word processing reported here, it appears that at least sometimes the processing time devoted to phrases might be a more appropriate metric. There is a wealth of literature where phrasal processing difficulty has been considered the appropriate metric in the investigation of syntactic parsing processes (e.g., Frazier & Rayner, 1982; Van Gompel, Pickering & Traxer, 2001; see also Murray, 2000), but work associated with testing models of 'eye movement control during *reading'* has been almost entirely lexico-centric.

But while orthographic preview effects were observed on words n, n+1, n+2 and the in spillover region, plausibility-related preview effects tended to be localised to word n+1. There was some evidence for a plausibility-related word n+1 parafoveal-on-foveal effect, but the effect size was small and as such, it would benefit from replication. And while there was some evidence that an anomalous word n+2 preview resulted in later first landing positions on word n+1 when n+1 was invalid, the qualifying interaction was only marginally significant, and so again, this effect would benefit from replication. It is noteworthy that the same pattern of plausibility-related preview effects were apparent on word n+2 as found on word n+1. However, these effects were clearly non-significant despite best attempts to create optimal conditions for word n+2 pre-processing. Thus, while orthographic preview effects appear to be distributed across several words, plausibility-related preview effects seem to be localised more closely to word n+1, and do not appear to stretch to word n+2. Presumably, this pattern reflects a difference in the stage of processing, with plausibility-related effects dependent on later lexical or post-lexical processing.

As discussed in Chapter 5, it is unlikely that the sole cause of preview effects can be attributed to the 'benefit' of having the parafoveal word available for inspection prior to direct fixation. Rather, a change from preview to target is also likely to produce a degree of interference, and this interference is also likely to increase inspection times on the target word (Murray et al, 2013; Risse & Kliegl, 2013). This hypothesis appears to be capable of explaining the pattern of effects reported in Chapter 5, where an illegal nonword preview produced a higher cost than real word previews when in the position of word n+1 but not n+2. A higher degree of confidence in the peculiarity might attract attention when the illegal nonword is in the position of word n+1, which could act to increase interference on the target word when it is fixated. When the nonword preview was located on word n+2, where acuity constraints might produce a lower degree of confidence in the peculiarity, it appeared to produce less interference upon direct fixation.

Perhaps the strongest evidence that preview effects stem, in part at least, from interference stemming from a parafoveal word change can be seen in Experiment 4. Here there was evidence that an anomalous preview resulted in increased target word inspection times compared to when the preview had been plausible. If preview effects were simply caused by the 'benefit' of having a word parafoveally available for pre-processing prior to fixation, then these differences should not have arisen. Rather, the pattern of effects in Experiment 4 suggest that the anomalous preview had been parafoveally processed, and this produced an increased degree of interference compared with when the preview had been plausible. This interpretation also appears to be consistent with the results of a variety of 'semantic preview benefit' experiments, such as those recently reported by Rayner and Schotter (2014).

It is possible, therefore, that the distribution of preview effects is determined by two factors (a) the level of processing that has been achieved on the parafoveal word, and (b) whether or not the nature of the preview attracted attention. Models of word recognition suggest that orthographic information is extracted prior to semantic information (e.g., Coltheart et al, 2001). Therefore, for higher-level preview effects to arise, more time will be required for the processing of the parafoveal word compared to that required for orthographic preview effects. Consequently, higher-level pre-processing is most likely to be restricted to immediately adjacent words. Manipulations related to word meaning might also be less likely to attract attention, since they maintain orthographic regularity, while illegal nonword previews do not. If such modulation in the breadth of attention is possible, then less attention means less parafoveal processing and, consequently, less interference upon direct fixation. Accordingly, distributed effects are less likely to arise when the preview manipulation employs real-word as opposed to illegal nonword previews.

Conversely, illegal nonword previews are more likely to attract attention, allowing a wider distribution of interference effects. As just discussed, however, the degree to which these effects cause interference might be offset against the reader’s confidence in the peculiarity. Thus, at more remote sites, such as word n+2, there might be higher confidence in perceiving an anomalous preview, since this is necessarily tied to the word recognition process, but a lower confidence in a nonword preview which has not yet fed into lexical processing. It appears therefore that preview effects result from a complex interaction between at least two (and potentially many more) factors, that make their expression unpredictable. That these effects arise at all, however, suggest that a loose coupling between fixation location and attention exists, with the results of the experiments reported here suggesting that the detachment might stretch as far as three word units. Again, this appears to question the validity of basing eye movement control models on individual word units, rather than phrases and calls into question the presupposition of the serial perspective that the breadth of attention might be variable – with word length – but never extend beyond the word boundary.

7.4. Accounting for Seemingly Parallel Effects within a Serial Framework

At first approximation, the level of distributed processing reported in this thesis appears to be in conflict with the idea that lexical processing proceeds in a strictly serial sequential fashion. As stated above, however, models of eye movement control have allowed an insight into how some seemingly parallel effects might be accounted for within a serial framework. This section is dedicated to assessing the plausibility of a series of auxiliary assumptions that have been factored into the E-Z Reader model in order to allow it to account for some of the effects reported here.

7.4.1. Distributed Processing and the Low Level Attentional Scan

The idea that a low level attentional scan could be responsible for orthographic parafoveal-on-foveal effects or word n+2 preview effects was first introduced into the E-Z Reader model in its 7th version (Reichle et al, 2003). While not currently instantiated in the model, it is argued that reading strategies could be altered based upon the detection of peculiarity in the parafovea. Such an assumption, therefore, potentially provides an explanation for orthographic parafoveal-on-foveal trends and effects reported in Experiments 2 and 3. Given that the E-Z Reader model is made up of separate modules, the addition of a low level attentional scanning module should be relatively easy to implement. It will be interesting to see how such a module modulates the predicted eye movement record.

The results of this thesis, however, call into question whether a low level attentional scan could really account for some of the effects related to the word n+2 preview manipulations. The idea that this might be the case was proposed by Angele and Rayner (2011), who suggested that a low level attentional scan might be implicated in modulating word n+1 ‘where’ decisions. Like Angele and Rayner, the present thesis reports evidence that word n+2 preview modulated skipping rate on word n+1, although the origin of these modulations appear to be routed in word n skipping probability. Thus, for the low level attentional scan to produce the patterns of effects reported here, one must assume a very broad range that encompasses at least 3-words.

It should be noted, however, that such a mechanism appears to fall short of explaining why sometimes parafoveal illegality appears to attract attention (e.g., encouraging word n+1 to be skipped), while at other times, it appears to encourage a more cautious reading strategy to be adopted (increasing fixation probability on word n+1). Including a low level attentional scan module in the model is therefore important to determine whether or not such a range of effects can indeed be modelled.

Finally, the low level attentional scan cannot explain word n+2 preview effects that do not involve orthographically illegal previews (e.g., Experiments 5 and 6). Overall, therefore, a low level attentional scan might be responsible for some orthographic parafoveal-on-foveal effects. It does not, however, readily explain some of the word n+2 preview effects reported here. But as Rayner (2009) suggests, this should be tested quantitatively before firm conclusions are drawn.

7.4.2. Fast Successive Parafoveal Lexical Processing

As will be recalled from Chapter 1, because attention shifts are decoupled from saccadic programming in the E-Z Reader model, attention can precede an overt eye movement. This decoupling typically only allows for a small lag before the eye is reunited with attention and is most prevalent when the fixated word is relatively ‘easy’ to process. Given that, in these studies, word n was always a high frequency word of medium length, and word n+1 was also always high frequency, it could be suggested that the experiments reported here provide the perfect conditions for a quick succession of lexical processing to advance prior to the eye leaving the fixated word.

Recent simulations with the E-Z Reader model have shown that word n+2 preview effects can be accounted for by a double attention shift – associated with a fast succession of lexical processing prior to the eye leaving the foveal word. While this explanation appears to hold for orthographic preview effects, the required amount of decoupling appears too great to allow for word n+2 preview effects that are semantic in nature (Schotter et al, 2014). A fast succession of lexical processing therefore appears to, given current instantiations in the model of the degree of flexibility between fixation location and the locus of attention, unable to explain the patterns of plausibility related preview effects reported in Experiments 5 and 6.

Furthermore, for a double-attention shift to take place within a serial model, such as the E-Z Reader, a necessary prerequisite is that the intervening word n+1 can be processed. Without this condition being met, lexical processing should halt on word n+1, preventing attention being shifted to the processing of word n+2. Word n+2 preview effects were, however, reported in Experiments 3 and 5 when word n+1 had received an invalid preview, suggesting that a double attention shift of this sort does not always occur when word n+ 2 preview effects arise.

A fast succession of lexical processing, it has been argued, could also account for word n+1 preview effects that appear to reflect the engagement of higher level processing on the parafoveal word (e.g., Experiment 4). Again, recent simulations using the E-Z Reader model appear to corroborate this suggestion. However, as will be recalled from Chapter 5, further simulations are needed to determine whether or not the model is able to simulate these effects contingent upon word n+1 being fixated (a necessary prerequisite for word n+1 semantic preview effects to arise). Also, it remains to be determined whether or not there is enough slack in the model to allow for parafoveal word processing to advance beyond semantic extraction and to allow the identified word to undergo some higher-level interpretation, as the results of Experiment 4 would seem to suggest. Based on the simulations of Schotter et al, it seems unlikely that the model could cope with this level of decoupling in its current form.

7.4.3. Distributed Processing and Mislocated Fixations

As discussed in detail in Chapters 4, 5 and 6, the proponents of the E-Z Reader model have invoked the concept of a mislocated fixation followed by a stay and process response to account for lexical parafoveal-on-foveal effects (such as those reported in Experiments 2 and 4) and semantic preview effects (e.g. Experiment 4). The validity of this explanation was tested in Experiment 7 by simulating mislocated fixations and measuring the immediate impact on lexical processing. This experiment provided evidence for an asymmetrical response to mislocated fixations. Critically, there was no evidence that a simulated undershoot resulted in a stay and process response. Quite the contrary – the results obtained here suggest that readers typically made a quick error-correcting saccade in response to undershooting one's target.

Since the results of Experiment 7 appear to suggest that saccadic overshoots result in a stay and process response, it could be the case that overshooting one's target word might result in the sort of orthographic spillover effects that have been reported throughout this thesis. For example, if word n+1 had been targeted, but saccadic error had resulted in it being skipped, this would permit word n+1 to be processed while fixating word n+2, thus producing an orthographic spillover effect. It should be noted, however, that while such a scenario is possible, given that the orthographic spillover effect of Experiment 5 persisted following a fixation of word n+1, this response would presumably need to be engaged followed a failed refixation of word n+1. Given that word n+1 was always a high frequency word that ranged between 3- and 6-letters in the present sets of experiments, it is seems rather unlikely that word n+1 would require a refixation frequently enough to be driving the spillover effects observed on word n+2, especially considering that only those refixations that overshoot can be driving the effect. Rather, it seems more likely that a fixation aimed at the following word can be programmed before the lexical processing of the current word is sufficiently complete for all orthographic effects to have been resolved.

Therefore, while a systematic response to a mislocated fixation could possibly impact the eye movement record, potentially producing the sorts of orthographic spillover effects reported here, there does not appear to be any evidence compatible with the view that an undershoot followed by a stay and process response is responsible for either lexical parafoveal-on-foveal effects or semantic preview effects.

7.5. Methodological Complications Associated with Investigating Preview Effects

Investigations into preview effects have typically employed the gaze contingent display change paradigm. This paradigm has provided an invaluable window into investigating which levels of representation are typically extracted from parafoveal words prior to fixation. This research suggests that the very early stages of word processing (e.g., orthographic and phonological extraction) have time to accrue on a parafoveal word prior to fixation (e.g., Rayner, 1975 and Pollatsek, Lesch, Morris & Rayner, 1992, respectively). However, until very recently, the possibility that ‘preview benefits’ might be confounded with interference effects (stemming from the word change) has not been fully appreciated, and it is possible that this has hindered the investigation of semantic preview effects. Emerging evidence from studies investigating potential interference effects in gaze contingent display change experiments (e.g. Murray et al, 2013; Risse & Kliegl, 2013) align with both the results of Schotter (2013) and those reported here in Experiment 4, in suggesting that previous failed attempts at uncovering higher level processing on parafoveal words may have been caused by information derived from before the word change interfering with, rather than facilitating, target word processing.

A second issue with using the gaze contingent display change paradigm to investigate preview effects is related to experiments investigating word n+2 pre-processing. This variation of the paradigm was first introduced by Rayner et al (2004) and, at first approximation, would appear to provide an insight into the range over which preview effects occur during reading. However, quite unlike the situation when the boundary is passed in word n+1 preview experiments, in word n+2 preview experiments, word n+2 will often be presented in its valid form prior to direct fixation. Indeed, this is necessarily the case unless the intervening word n+1 is skipped. Thus, not only are word n+2 preview effects expected to be both numerically and statistically smaller than those occurring on word n+1 due to greater acuity constraints, but information derived about word n+2 during the fixation of word n+1 might further ‘wash out’ any marginal effect that might have accrued. It seems therefore, that gaze contingent display change experiments which set out to investigate word n+2 preview effects are likely to be biased towards the null hypothesis. Indeed, this might explain why the same pattern of word n+2 plausibility-related preview effects were present on word n+2 (Experiment 5) as were present on word n+1 (Experiment 4), but failed to achieve significance.

7.6. Parafoveal Preview Effects and Parallel Models of Eye Movement Control

This thesis has provided evidence for orthographic and lexical parafoveal-on-foveal effects, orthographic and higher-level preview effects stemming from words n+1 and n+2, and orthographic spillover effects. All these results appear to fit most parsimoniously with a model of eye movement control, such as SWIFT, that can allow for the processing of multiple words simultaneously. Furthermore, the results of Experiment 1 appears to support the proposal that multiple words can be lexically processed in a simultaneous fashion. However, several results in this thesis do not necessarily sit as comfortably as one might expect with the SWIFT model.

First, it is worth noting that, unlike the E-Z Reader model, the predictions of the SWIFT model are often difficult to determine. The complex nature of the model makes it difficult even for its own architects to fully understand how it can account for some effects. For example, the model's ability to correctly simulate both skipping costs and benefits (following the incorporation of the zoom lens mechanism) was a surprise to Schad and Engbert (2012), with an explanation for why this might be the case needing to be deferred until further explorations with the model have been completed. And while the model can accommodate parafoveal-on-foveal effects, which it does via modulations in fixation and refixation probabilities of all words falling within the effective span of apprehension, it is not known, and indeed, cannot be easily predicted by a non-modelling expert, whether it could cope with unorthodox as well as orthodox parafoveal-on-foveal effects. Indeed, the model's opaque nature is one of the major criticisms typically levelled at the SWIFT model (e.g. Reichle et al, 2009).

Also, while the model has recently simulated word n+2 preview effects (Risse, Hohenstein, Kliegl & Engbert, 2014), with longer inspection time on word n+2 when its activation level was reset to zero after passing the invisible boundary, it is unclear whether it could account for word n+3 preview effects which act to modify word n and word n+1 skipping probabilities. Even assuming that the span of attention is fully dilated, it seems somewhat implausible to assume that a word 3-words downstream should be able to achieve enough activation to influence skipping probabilities on words n and n+1. Again, further simulations will be required to determine whether or not the model can cope with such a broad effective range of attention.

The findings reported here (Experiment 2) and by Kliegl et al (2007), showing that alternative high frequency words appear to facilitate word n+2 pre-processing more successfully than a determiner, is also not easy to explain within the architecture of the SWIFT model. As discussed above, the most plausible explanation for these effects is that the reader is more likely to seek information from word n+2 when the complete interpretation of word n+1 depends upon the subsequent word's meaning. However, despite being a parallel model, SWIFT does not contain a mechanism to account for higher-level integrative effects. Indeed, like the E-Z Reader model, SWIFT bases its modelling efforts at the individual word level, rather than at the phrasal level, which is perhaps surprising given that it promotes the idea that multiple words are processed simultaneously. Of course, not accounting for higher-level processing effects also means that the model is currently unable to accommodate the plausibility-related preview effects reported in Experiment 4. While this is not an insurmountable problem for the model, it will be interesting to see how successfully it can evolve to accommodate such higher-level effects.

One final finding that appears to be problematic for the SWIFT model comes from Experiment 7. Here there was evidence that both within- and between-word undershoots caused the reader to initiate a quick error correcting saccade. The same was not true, however, for saccadic overshoots, where there was no evidence for initiation of a quick error correction. This is contrary to the assumptions of the SWIFT model, which assumes that all fixations falling on an erroneous word - and only on an erroneous word - trigger a quick error-correcting saccade. Further, there was no evidence either in this experiment, or in that of Feng (2011), to suggest that the mislocated fixations were detected via an efference copy, as the proponents of SWIFT suggest. Remodelling SWIFT to account for these differences will undoubtedly have an impact on the simulation output of the model, whether this might enhance the model, or indeed be to its detriment remains to be determined.

Overall, therefore, the results presented in this thesis appear to pose a problem even for a model that assumes that the reading process can advance using parallel lexical processing. While the results appear to fit most parsimoniously with a parallel model, it is clear that SWIFT is currently too underspecified to account for the complexity of the findings obtained in the set of studies reported in this thesis.

7.7. Distinguishing Parallel and Serial Models of Eye Movement Control

As described above, the core architecture of the E-Z Reader model allows for a degree of decoupling between fixation location and attention, allowing it to account for effects such as semantic preview effects and spillover effects. Additional auxiliary assumptions have also been added, such as the low level attentional scan. These might allow it to simulate results that are more akin to what would be expected from a parallel model. Likewise, the SWIFT model can mimic a pattern of eye movements that one might predict based on a serial architecture, especially when the text is difficult and the zoom lens remains contracted. Despite the fundamental differences in the core architectures of these models, therefore, the predictions of both can align, so that distinguishing between the two models becomes increasing difficult.

This conundrum has recently been raised by Murray et al (2013). They propose that the way forward might now be restricted to assessing the plausibility of the auxiliary assumptions that the models have reverted to in order to explain some effects. Chapter 7 of this thesis has attempted to take the research in this direction by testing a key, but principally untested, auxiliary assumption of the E-Z Reader model. These results call into question the plausibility of the mislocated fixation hypothesis in accounting for lexical parafoveal-on-foveal effects, and as a consequence, now challenges the proponents of the E-Z Reader model to find an alternative explanation for parafoveal-on-foveal effects.

Of course, that is not to say that it is pointless to continue investigating the distribution of attention during reading. Indeed, these investigations have proved invaluable for challenging the models and encouraging their evolution. It is simply suggested that more attention needs to be focussed on experimentally testing auxiliary assumptions rather than relying on qualitative descriptions. In many ways, this mirrors Rayner’s (2009) sentiment that nothing should be assumed regarding what models can and cannot account for. As researchers, it is imperative that we fully test these assumptions experimentally *and* with appropriate simulations.

7.8. Conclusions

This thesis has reported a variety of effects that appear to challenge both the E-Z Reader and SWIFT models. These effects include orthographic and lexical parafoveal-on-foveal effects, plausibility-related word n+1 preview effects and a series of word n+2 preview trends and effects. Hopefully it is clear from the discussions above, that while the SWIFT model is currently too underspecified to accommodate many of these findings, the E-Z Reader model can only do so, at present, via verbal explanation associated with various axillary assumptions. What seems very clear from the results reported here, however, is that there must be a degree of decoupling between fixation location and attention, with some evidence suggesting that this might, under some circumstances, extend up to 3-word units. But while this might initially seem implausible, there is some other evidence in the literatures (e.g. Kennedy, Murray, Boissiere, 2004) indicating the presence of such long-range effects.

Experiment 1 of this thesis has provided evidence which suggests that the lexical processing of multiple words is, at the very least, possible. It does not, however, show that this is necessarily the default process. The patterns of results presented here are, nevertheless, most consistent with the idea that multiple words can be processed in an overlapping fashion. This is the case, not least of all, because the results of Experiment 7 suggest that the mechanism suggested by the proponents of E-Z Reader to accommodate lexical parafoveal-on-foveal effects is not at all likely.

References

Aaronson, D., & Scarborough, H. S. (1976). Performance Theories for Sentence Coding: Some Quantitative Evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(1), 56–70.

Adams, M. W. J., Wood, D., & Carpenter, R. H. S. (2000). Expectation acuity: the spatial specificity of the effect of prior probability on saccade latency. *Journal of Physiology*, *527*(Suppl), 140–141.

Altarriba, J., Kambe, G., Pollatsek, A., & Rayner, K. (2001). Semantic codes are not used in integrating information across eye fixations in reading: evidence from fluent Spanish-English bilinguals. *Perception & Psychophysics*, *63*(5), 875–90.

Angele, B., & Rayner, K. (2011). Parafoveal processing of word n + 2 during reading: do the preceding words matter? *Journal of Experimental Psychology: Human Perception and Performance*, *37*(4), 1210–20.

Angele, B., & Rayner, K. (2013). Processing the in the parafovea: are articles skipped automatically? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 649–62.

Angele, B., Slattery, T. J., Yang, J., Kliegl, R., & Rayner, K. (2008). Parafoveal processing in reading: Manipulating n+1 and n+2 previews simultaneously. *Visual Cognition*, *16*(6), 697–707.

Ashby, J., & Rayner, K. (2004). Representing syllable information during silent reading: Evidence from eye movements. *Language and Cognitive Processes*, *19*(3), 391–426.

Ashby, J., Treiman, R., Kessler, B., & Rayner, K. (2006). Vowel processing during silent reading: evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 416–24.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Baccino, T., & Manunta, Y. (2005). Eye-Fixation-Related Potentials : Insight into Parafoveal Processing. *Journal of Psychophysiology*, *19*(3), 204–215.

Balota, D. A. (1990). The role of meaning in word recognition. In D. A. Balota, G. B. Flores d’Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 9–32). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, *17*(3), 364–390.

Balota, D. A., & Rayner, K. (1983). Parafoveal visual information and semantic contextual constraints. *Journal of Experimental Psychology: Human Perception and Performance*, *9*(5), 726–38.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Barron, J. (1975). Successive stages in word recognition. In P.M.A.Rabbit & S. Dorne (Eds.), *Attention and performance V*. New York: Academic Press.

Barron, R. W., & Henderson, L. (1977). The effects of lexical and semantic information on same-different visual comparison of words. *Memory & Cognition*, *5*(5), 566–79.

Barron, R. W., & Pittenger, J. B. (1974). The effect of orthographic structure and lexical meaning on “same-different” judgments. *Quarterly Journal of Experimental Psychology*, *26*(4), 566–581.

Bates D, Maechler M, Bolker B and Walker S (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.0-5, http://CRAN.R-project.org/package=lme4.

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research*, *19*(9), 967–983.

Bicknell, K., Higgins, E., Levy, R., & Rayner, K. (2013). Evidence for cognitively controlled saccade targeting in reading. In *Proceedings of the 35th annual conference of the Cognitive Science Society* (Vol. 1, pp. 197–202).

Brysbaert, M., Drieghe, D., & Vitu, F. (2005). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (pp. 53–78). Oxford, UK: Oxford University Press.

Calvo, M. G., & Castillo, M. D. (2005). Processing of Threat-related Information Outside the Focus of Visual Attention. *The Spanish Journal of Psychology*, *8*(1), 3–11.

Carpenter, P. A., & Daneman, M. (1981). Lexical Retrieval and error recovery in reading: A model based on eye fixations. *Journal of Verbal Learning and Verbal Behavior*, *20*(2), 137–160.

Carpenter, P. A., & Just, M. A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 275–307). New York: Academic Press.

Chambers, S. M., & Forster, K. I. (1975). Evidence for lexical access in a simultaneous matching task. *Memory & Cognition*, *3*(5), 549–59.

Chung, S. T., Mansfield, J. S., & Legge, G. E. (1998). Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vision Research*, *38*(19), 2949–62.

Clahsen, H., Hong, U., & Sonnenstuhl-Henning, I. (1995). Grammatical constraints in syntactic processing: Sentencematching experiments on German. *The Linguistic Review*, *12*, 5–33.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–56.

Deubel, H., Regan, J. K. O., & Radach, R. (2000). Attention , information processing , and eye movement control. In A. Kennedy, R. . Radach, D. Heller, & J. Pynte (Eds.), *Reading as a Perceptual Process* (Elsevier., pp. 355–374). Oxford: Elsevier

Drieghe, D. (2011). Parafoveal-on-foveal effects on eye movements during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 839–855). New York: Oxford University Press.

Drieghe, D., Brysbaert, M., & Desmet, T. (2005). Parafoveal-on-foveal effects on eye movements in text reading: does an extra space make a difference? *Vision Research*, *45*(13), 1693–706.

Drieghe, D., Brysbaert, M., Desmet, T., & De Baecke, C. (2004). Word skipping in reading: On the interplay of linguistic and visual factors. *European Journal of Cognitive Psychology*, *16*(1-2), 79–103.

Drieghe, D., Pollatsek, A., Staub, A., & Rayner, K. (2008). The word grouping hypothesis and eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1552–60.

Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(5), 954–9.

Drieghe, D., Rayner, K., & Pollatsek, A. (2008). Mislocated fixations can account for parafoveal-on-foveal effects in eye movements during reading. *Quarterly Journal of Experimental Psychology*, *61*(8), 1239–49.

Engbert, R., & Kliegl, R. (2011). Parallel graded attention models of reading. In S. P. Liversedge, I. . Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 287–800). Oxford: Oxford University Press.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, *42*(5), 621–36.

Engbert, R., Nuthmann, A., & Kliegl, R. (2007). An iterative algorithm for the estimation of the distribution of mislocated fixations during reading. In R.

P. G. Van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye Movements: A Window on Mind and brain* (pp. 319–337). Oxford: Elsevier.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: a dynamical model of saccade generation during reading. *Psychological Review*, *112*(4), 777–813.

Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, *40*(4), 225–240.

Erlhagen, W., & Schöner, G. (2002). Dynamic field theory of movement preparation. *Psychological Review*, *109*(3), 545–572.

Feng, G. (2009). Mixed Responses : Why Readers Spend Less Time at Unfavorable Landing Positions. *Journal of Eye Movement Research*, *3*(2), 1–26.

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-Enough Representations in Language Comprehension. *Current Directions in Psychological Science*, *11*(1), 11–15.

Ferreira, F., & Patson, N. D. (2007). The “Good Enough” Approach to Language Comprehension. *Language and Linguistics Compass*, *1*(1-2), 71–83.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Forster, K. I. (1976). Accessing the Mental Lexicon. In R. J. Wales & E. C. T. Walker (Eds.), *New Approaches to Language Mechanisms*. Amsterdam: North Holland.

Foxe, J. J., & Simpson, G. V. (2002). Flow of activation from V1 to frontal cortex in humans. A framework for defining “early” visual processing. *Experimental Brain Research*, *142*(1), 139–50.

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: lexicon and grammar*. Boston: Houghton Mifflin.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*(2), 178–210.

Freedman, S. E., & Forster, K. I. (1985). The psychological status of overgenerated sentences. *Cognition*, *19*(2), 101–131.

Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(5), 862–77.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(3), 417–29.

Henderson, J. M., & Ferreira, F. (1993). Eye movement control during reading: fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale*, *47*(2), 201–21.

Henderson, L. (1980). Is there a lexicality component in the word superiority effect? *Perception & Psychophysics*, *28*(2), 179–184.

Hernández-Peón, R. (1964). Psychiatric implications of neurophysiological research. *Bulletin of the Menninger Clinic*, *24*, 165–185.

Hogaboam, T. W. (1983). Reading patterns in eye movements. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 309–332). New York: Academic Press.

Hohenstein, S., & Kliegl, R. (2013). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 166–90.

Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic preview benefit in eye movements during reading: A parafoveal fast-priming study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1150–70.

Howes, D. H., & Solomon, R. L. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology*, *41*(6), 401–410.

Hyönä, J., & Bertram, R. (2004). Do frequency characteristics of nonfixated words influence the processing of fixated words during reading? *European Journal of Cognitive Psychology*, *16*(1-2), 104–127.

Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye-movement-contingent display change study. *Memory & Cognition*, *32*(4), 523–32.

Hyönä, J., & Häikiö, T. (2005). Is emotional content obtained from parafoveal words during reading? An eye movement analysis. *Scandinavian Journal of Psychology*, *46*, 475–83.

Inhoff, A. W. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning and Verbal Behavior*, *23*(5), 612–624.

Inhoff, A. W. (1989). Lexical access during eye fixations in reading: Are word access codes used to integrate lexical information across interword fixations? *Journal of Memory and Language*, *28*(4), 444–461.

Inhoff, A. W., Radach, R., & Eiter, B. (2006). Temporal overlap in the linguistic processing of successive words in reading: reply to Pollatsek, Reichle, and

Rayner (2006a). *Journal of Experimental Psychology. Human Perception and Performance*, *32*(6), 1490–5.

Inhoff, A. W., Radach, R., Starr, M., & Greenberg, S. (2000). Allocation of visuo-spatial attention and saccade programming during reading. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a Perceptual Process* (pp. 221–246). Oxford: Elsevier.

Inhoff, A. W., Starr, M., & Shindler, K. L. (2000). Is the processing of words during eye fixations in reading strictly serial? *Perception & Psychophysics*, *62*(7), 1474–84.

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., Gathercole, S. E., & Rayner, K. (2008). Children and adults' processing of anomaly and implausibility during reading: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, *61*(5), 708–723.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review*, *87*, 239–254.

Kalesnykas, R. P., & Hallett, P. E. (1994). Retinal eccentricity and the latency of eye saccades. *Vision Research*, *34*(4), 517–531.

Kapoula, Z. (1985). Evidence for a range effect in the saccadic system. *Vision Research*, *25*(8), 1155–1157.

Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing trade- off. In G. Underwoord (Ed.), *Eye guidance in reading and scene perception* (pp. 149–223). Oxford: Elsevier.

Kennedy, A. (2000). Attention allocation in reading. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a Perceptual Process* (pp. 193–220). Oxford, UK: Elsevier.

Kennedy, A. (2003). The Dundee Corpus [CD-ROM]. *Psychology Department, University of Dundee*.

Kennedy, A., Murray, W., & Boissiere, C. (2004). Parafoveal pragmatics revisited. *European Journal of Cognitive Psychology*, *16*(1-2), 128–153.

Kennedy, A., & Murray, W. S. (1987). Spatial coordinates and reading: Comments on Monk (1985). *The Quarterly Journal of Experimental Psychology Section A*, *39*(4), 649–656.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, *45*(2), 153–68.

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *The Quarterly Journal of Experimental Psychology. Section A: Human Experimental Psychology*, *55*(4), 1307–37.

Kliegl, R. (2007). Toward a perceptual-span theory of distributed processing in reading: A reply to Rayner, Pollatsek, Drieghe, Slattery, and Reichle (2007). *Journal of Experimental Psychology: General*, *136*(3), 530–537.

Kliegl, R., & Engbert, R. (2003). SWIFT Explorations. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 391–412). Amsterdam: Elsevier.

Kliegl, R., & Engbert, R. (2005). Fixation durations before word skipping in reading. *Psychological Bulletin & Review*, *12*(1), 132–138.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*(1-2), 262–284.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: the influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology. General*, *135*(1), 12–35.

Kliegl, R., Risse, S., & Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects from word n + 2. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(5), 1250–5.

Laubrock, J., & Hohenstein, S. (2012). Orthographic consistency and parafoveal preview benefit: A resource-sharing account of language differences in processing of phonological and semantic codes. *The Behavioral and Brain Sciences*, *35*(5), 292–293.

Legge, G. E., Mansfield, J. S., & Chung, S. T. (2001). Psychophysics of reading. XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision Research*, *41*(6), 725–43.

Liversedge, S. P., Rayner, K., White, S. J., Vergilino-Perez, D., Findlay, J. M., & Kentridge, R. W. (2004). Eye movements when reading disappearing text: is there a gap effect in reading? *Vision Research*, *44*(10), 1013–24.

Matin, E. (1974). Saccadic Suppression: A Review. *Psychological Bulletin*, *81*, 899–917.

McConkie, G., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, *17*, 578–586.

McConkie, G., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society*, *8*, 365–368.

McConkie, G. W. (1979). On the role and control of eye movements in reading. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), *Processing of visible language: Volume 1* (pp. 37–48). New York: Plenum Press.

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations on words. *Vision Research*, *28*(10), 1107–1118.

Mcconkie, G. W., Kerr, P. W., Reddix, M. D., Zola, D., & Jacobs, A. M. (1989). Eye movement control during reading: II. Frequency of refixating a word. *Perception & Psychophysics*, *46*(3), 245–253.

Mcconkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading? *Perception & Psychophysics*, *25*(3), 221–224.

McDonald, S. A. (2005). Parafoveal preview benefit in reading is not cumulative across multiple saccades. *Vision Research*, *45*(14), 1829–34.

Meuter, R. F. I., & Allport, A. (1999). Bilingual Language Switching in Naming : Asymmetrical Costs of Language Selection, *40*, 25–40.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227–234.

Miellet, S., O’Donnell, P. J., & Sereno, S. C. (2009). Parafoveal magnification: visual acuity does not modulate the perceptual span in reading. *Psychological Science*, *20*(6), 721–8.

Miellet, S., & Sparrow, L. (2004). Phonological codes are assembled before word fixation: Evidence from boundary paradigm in sentence reading. *Brain and Language*, *90*(1-3), 299–310.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance,*, *10*(5), 667–682.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, *76*(2), 165–178.

Mouchetant-Rostaing, Y., Giard, M. H., Bentin, S., Aguera, P. E., & Pernier, J. (2000). Neurophysiological correlates of face gender processing in humans. *The European Journal of Neuroscience*, *12*(1), 303–10.

Murray, W. S. (1982). *Sentence matching: The influence of meaning and structure. Unpublished doctoral dissertation*. Monash University, Victoria, Australia.

Murray, W.S. (1998). Parafoveal pragmatics. In G. Underwood (Ed.). *Eye Guidance in Reading and Scene Perception*. (pp.181-199) Oxford: Elsevier.

Murray, W. S. (2000). Sentence processing: Issues and measures. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a Perceptual Process* (pp. 649–664). Oxford: Elsevier.

Murray, W.S. (2006). The nature and time course of pragmatic plausibility effects. *Journal of Psycholinguistic Research (Special Issue)*, 35(1), 79-99.

Murray, W. S., Fischer, M. H., & Tatler, B. W. (2013). Serial and parallel processes in eye movement control : Current controversies and future directions. *The Quarterly Journal of Experimental Psychology*, *66*(3), 417–428.

Murray, W. S., & Forster, K. I. (2004). Serial mechanisms in lexical access: the rank hypothesis. *Psychological Review*, *111*(3), 721–56.

Murray, W. S., & Forster, K. I. (2008). The rank hypothesis and lexical decision: a reply to Adelman and Brown (2008). *Psychological Review*, *115*(1), 240–52.

Murray, W. S., Rayner, K., & Wakeford, L. J. (2013). Preview benefit or preview cost? In *17th European Conference on Eye Movements*. Lund, Sweden.

Murray, W. S., & Rowan, M. (1998). Early, Mandatory, Pragmatic Processing. *Journal of Psycholinguistic Research*, *27*(1), 1–22.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. *Basic Processes in Reading: Visual Word Recognition*, *11*, 264–336.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision Research*, *45*(17), 2201–17.

Nuthmann, A., Engbert, R., & Kliegl, R. (2007). The IOVP effect in mindless reading: Experiment and modeling. *Vision Research*, *47*(7), 990–1002.

O'Regan, J. K., & Jacobs, A. M. (1992). Optimal Viewing Position Effect in Word Recognition : A Challenge to Current Theory. *Journal of Experimental Psychology: Human Perception and Performance.*, *18*(1), 185–197.

O'Regan, J. K., Lévy-Schoen, A., Pynte, J., & Brugaillère, B. (1984). Convenient fixation location within isolated words of different length and structure. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(2), 250–7.

O'Regan, K. (1979). Saccade size control in reading: Evidence for the linguistic control hypothesis. *Perception & Psychophysics*, *25*(6), 501–509.

O'Regan, K. (1981). The convenient viewing position hypothesis. In D. L. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye Movements, Cognition and visual perception* (pp. 289–298). Hillside, NJ: Erlbaum.

Pollatsek, A., & Digman, L. (1977). Dependent spatial channels in visual processing. *Cognitive Psychology*, *9*(3), 326–352.

Pollatsek, A., Juhasz, B. J., Reichle, E. D., Machacek, D., & Rayner, K. (2008). Immediate and delayed effects of word frequency and word length on eye movements in reading: a reversed delayed effect of word length. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(3), 726–50.

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(1), 148–62.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, *52*(1), 1–56.

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, *32*(1), 3–25.

Pynte, J., Kennedy, A., & Ducrot, S. (2004). The influence of parafoveal typographical errors on eye movements in reading. *European Journal of Cognitive Psychology*, *16*(1-2), 178–202.

R Development Core Team (2012): *R: A language environment for statistical computing [Computer software].* Vienna. Austria: R Foundation for Statistical Computing.

Radach, R., Glover, L., & Vorstius, C. (2007). Exploring the limits of spatially distributed processing - a new look at n+2 preview effects. In *14th European Conference on Eye Movements*. Potsdam, Germany.

Radach, R., & Heller, D. (2000). Relations between spatial and temporal aspects of eye movement control. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 165–192). Oxford: Elsevier.

Radach, R., Inhoff, A. W., Glover, L., & Vorstius, C. (2013). Contextual constraint and N + 2 preview effects in reading. *Quarterly Journal of Experimental Psychology*, *66*(3), 619–33.

Radach, R., & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for future research. *European Journal of Cognitive Psychology*, *16*(1-2), 3–26.

Radach, R., & Kennedy, A. (2013). Eye movements in reading: Some theoretical context. *Quarterly Journal of Experimental Psychology*, *66*(3), 429–52.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*(1), 65–81.

Rayner, K. (1978). Foveal and parafoveal cues in reading. In J. Requin (Ed.), *Foveal and parafoveal cues in reading* (pp. 149–162). Hillsdale: Erlbaum.

Rayner, K. (1979). Eye guidance in reading: fixation locations within words. *Perception*, *8*(1), 21–30.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422.

Rayner, K. (2009a). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457–506.

Rayner, K. (2009b). Eye Movements in Reading: Models and Data. *Journal of Eye Movement Research*, *2*(5), 1–10.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: implications for the E-Z Reader model. *Journal of Experimental Psychology. Human Perception and Performance*, *30*(4), 720–732.

Rayner, K., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology*, *40*(4), 473–83.

Rayner, K., Binder, K. S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: word predictability has little influence on initial landing positions in words. *Vision Research*, *41*(7), 943–54.

Rayner, K., & Duffy, S. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*(3), 191–201.

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: eye movements during reading and scanning are different. *Perception & Psychophysics*, *58*(5), 734–47.

Rayner, K., Fischer, M. H., & Pollatsek, A. (1998). Unspaced Text Interferes with Both Word Identification and Eye Movement Control. *Vision Research*, *38*(8), 1129–1144.

Rayner, K., Inhoff, A. W., Morrison, R. E., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of Foveal and Parafoveal Vision During Eye Fixations in Reading. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(1), 167–179.

Rayner, K., Juhasz, B. J., & Brown, S. J. (2007). Do readers obtain preview benefit from word N + 2? A test of serial attention shift versus distributed lexical processing models of eye movement control in reading. *Journal of*

Experimental Psychology: Human Perception and Performance, *33*(1), 230–45.

Rayner, K., Liversedge, S. P., White, S. J., & Vergilino-Perez, D. (2003). Reading Disappearing Text: Cognitive Control of Eye Movements. *Psychological Science*, *14*(4), 385–388.

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research*, *16*, 829–837.

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology*, *12*(2), 206–226.

Rayner, K., & Morrison, R. E. (1981). Eye movements and identifying words in parafoveal vision. *Bulletin of the Psychonomic Society*, *17*(3), 135–138.

Rayner, K., Pollatsek, A., Drieghe, D., Slattery, T. J., & Reichle, E. D. (2007). Tracking the mind during reading via eye movements: comments on Kliegl, Nuthmann, and Engbert (2006). *Journal of Experimental Psychology. General*, *136*(3), 520–9.

Rayner, K., & Schotter, E. R. (2014). Semantic Preview Benefit in Reading English: The Effect of Initial Letter Capitalization, *40*(4), 1617–1628.

Rayner, K., Schotter, E. R., & Drieghe, D. (2014). Lack of semantic parafoveal preview benefit in reading revisited. *Psychonomic Bulletin & Review*, *21*(4), 1067–72.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, *4(3-4)*, 121–149.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton Jr, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, *4*(3-4), SI21–SI49.

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: a comparison of two types of models. *Journal of Experimental Psychology. Human Perception and Performance*, *22*(5), 1188–200.

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 514–28.

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(6), 1290–301.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, *3*(4), 504–9.

Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception & Psychophysics*, *31*(6), 537–50.

Rayner, K., McConkie, G.W., & Ehrlich, S. (1978). Eye movements and integrating across fixations. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(4), 529–544.

Rayner, K., White, S.J., Kambe, G., Miller, B., & Liversedge, S. P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 213–234). Amsterdam: Elsevier.

Reichle, E. D. (2011). Serial-attenton models of reading. In S. P. Liversedge, D. Gilchrest, Iain, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 767–786). Oxford: Oxford University Press.

Reichle, E. D., & Drieghe, D. (2015). Using E-Z Reader to Examine the Consequences of Fixation-Location Measurement Error, *41*(1), 262–270.

Reichle, E. D., Liversedge, S. P., Pollatsek, A., & Rayner, K. (2009). Encoding multiple words simultaneously in reading is implausible. *Trends in Cognitive Sciences*, *13*(3), 115–119.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*(1), 125–157.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E-Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research*, *7*(1), 4–22.

Reichle, E. D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research*, *39*(26), 4403–11.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z reader model of eye-movement control in reading: comparisons to other models. *The Behavioral and Brain Sciences*, *26*(4), 445–476.

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, *16*(1), 1–21.

Reilly, R. G., & O'Regan, K. (1998). Eye movement control during reading: Simulation of some word-targeting strategies. *Vision Research*, *38*(2), 303–317.

Reilly, R. G., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research*, *7*, 34–55.

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the E-Z Reader model. *Psychological Science*, *17*(9), 742–6.

Richter, E. M., Engbert, R., & Kliegl, R. (2006). Current advances in SWIFT. *Cognitive Systems Research*, *7*(1), 23–33.

Risse, S., Hohenstein, S., Kliegl, R., & Engbert, R. (2014). A Theoretical Analysis of the Perceptual Span based on SWIFT Simulations of the n + 2 Boundary Paradigm. *Visual Cognition*, *22*(3), 283–308.

Risse, S., & Kliegl, R. (2011). Adult age differences in the perceptual span during reading. *Psychology and Aging*, *26*(2), 451–60.

Risse, S., & Kliegl, R. (2013). Half preview benefit, half delayed cost: Evidence from reading with gaze-contigent preview manipulation of word n+1. In *European Conference on Eye Movements*. Lund, Sweden.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior*, *10*(6), 645–657.

Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model. *Visual Cognition*, *20 (4-5)*, 391–421.

Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: word frequency effects and individual differences. *Memory & Cognition*, *26*(6), 1270–1281.

Schotter, E. R. (2013). Synonyms provide semantic preview benefit in English. *Journal of Memory and Language*, *69*(4), 619–633.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception & Psychophysics*, *74*(1), 5–35.

Schotter, E. R., Reichle, E. D., & Rayner, K. (2014). Rethinking parafoveal processing in reading : Serial- attention models can explain semantic preview benefit and N + 2 preview effects. *Visual Cognition*, *22*(3-4), 309–333.

Scougal, A., & Murray, W. S. (2009). Syntactic Context and Lexical Effectson Eye Movements during Reading. In *15th European Conference on Eye Movements (ECEM15)*. Southampton, England.

Sereno, S. C., & Rayner, K. (1992). Fast Priming During Eye Fixations in Reading. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(1), 173–184.

Simola, J., Holmqvist, K., & Lindgren, M. (2009). Right visual field advantage in parafoveal processing: evidence from eye-fixation-related potentials. *Brain and Language*, *111*(2), 101–13.

Sparks, D. L. (2002). The brainstem control of saccadic eye movements. *Nature Reviews. Neuroscience*, *3*, 952–64.

Starr, M., & Inhoff, A. (2004). Attention allocation to the right and left of a fixated word: Use of orthographic information from multiple words during reading. *European Journal of Cognitive Psychology*, *16*(1-2), 203–225.

Staub, A., Rayner, K., Pollatsek, A., Hyönä, J., & Majewski, H. (2007). The time course of plausibility effects on eye movements in reading: evidence from noun-noun compounds. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(6), 1162–9.

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition*, *36*(1), 201–216.

Thibadeau, R., Just, M. A., & Carpenter, P. A. (1982). A model of the time course and content of reading. *Cognitive Science*, *6*, 157–203.

Treisman, A., M, & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.

Underwood, G., Binns, A., & Walker, S. (2000). Attentional demands on the processing of neighbouring words. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a Perceptual Process* (pp. 247–268). Oxford: Elsevier Ltd.

Van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in Sentence Processing: Evidence against Current Constraint-Based and Two-Stage Models. *Journal of Memory and Language*, *45*(2), 225–258.

VanRullen, R., & Thorpe, S. J. (2001). The time course of visual processing: from early perception to decision-making. *Journal of Cognitive Neuroscience*, *13*(4), 454–61.

Vergilino, D., & Beauvillain, C. (2000). The planning of refixation saccades in reading. *Vision Research*, *40*(25), 3527–38.

Vitu, F., Brysbaert, M., & Lancelin, D. (2004). A test of parafoveal-on-foveal effects with pairs of orthographically related words. *European Journal of Cognitive Psychology*, *16*(1-2), 154–177.

Vitu, F., Mcconkie, G. W., Kerr, P., & O'Regan, J. K. (2001). Fixation location effects on fixation durations during reading : an inverted optimal viewing position effect. *Vision Research*, *41*, 3513–3533.

Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic Bulletin & Review*, *14*(4), 770–5.

Whaley, C. P. (1978). Word–nonword classification time. *Journal of Verbal Learning & Verbal Behavior*, *17*(2), 143–154.

White, S. J. (2008). Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(1), 205–23.

White, S. J., Bertram, R., & Hyönä, J. (2008). Semantic processing of previews within compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 988–93.

White, S. J., & Liversedge, S. P. (2006). Foveal processing difficulty does not modulate non-foveal orthographic influences on fixation positions. *Vision Research*, *46*(3), 426–37.

White, S. J., Rayner, K., & Liversedge, S. P. (2005). Eye movements and the modulation of parafoveal processing by foveal processing difficulty : A reexamination. *Psychonomic Bulletin & Review*, *12*(5), 891–896.

White, S., & Liversedge, S. (2004). Orthographic familiarity influences initial eye fixation positions in reading. *European Journal of Cognitive Psychology*, *16*(1-2), 52–78.

Wyman, D., & Steinman, R. M. (1973). Latency characteristics of small saccades. *Vision Research*, *13*(11), 2173–2175.

Yang, J., Wang, S., Xu, Y., & Rayner, K. (2009). Do Chinese readers obtain preview benefit from word n + 2? Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(4), 1192–204.

Appendix A

Participant Information Sheet Used in Experiment 1.

PARTICIPANT INFORMATION SHEET

HOW MANY WORDS CAN WE PROCESS AT ONCE?

INVITATION TO TAKE PART IN A RESEARCH STUDY
You are being asked to take part in a research study, which will investigate your breadth of attention during reading. I am a PhD student in the psychology department and my supervisor is Dr Wayne Murray.

PURPOSE OF THE RESEARCH STUDY
The purpose of the research is to see whether information from two words can be extracted at once. Two strings of letters will be presented at a time on a computer screen. These two strings may or may not form a word. The task is to decide whether the two strings of letters are the same or different - some pairs of letter strings will differ by 1 character while other pairs will be identical. After a very brief presentation, the two strings of letters will disappear from the screen, at which point a decision of same or different is required from you.

Participation in this research would benefit me, Laura Wakeford as it will contribute towards my thesis.

TIME COMMITMENT
The study will require 45-60 minutes to complete with one visit/session.

TERMINATION OF PARTICIPATION
You may decide to stop being a part of the research study at any time without explanation. There will be no penalty and you will still receive you course credits.

RISKS
There are no known risks for you in this study.

COST, REIMBURSEMENT AND COMPENSATION
Your participation in this study is voluntary. You will receive 3 course credits after completion of the testing.

CONFIDENTIALITY/ANONYMITY
The data we collect do not contain any personal information about you. No one will link the data you provided to your identity and name.

FOR FURTHER INFORMATION ABOUT THIS RESEARCH STUDY
Laura Wakeford will be glad to answer your questions about this study at any time.
If you want to find out about the final results of this study, you should contact Laura Wakeford using the following email: ljwakefodr@dundee.ac.uk.

The University Research Ethics Committee of the University of Dundee has reviewed and approved this research study.

UREC v. 1.9, 15 December 2006

Appendix B

Consent Form Used in Experiment 1.

INFORMED CONSENT FORM

HOW MANY WORDS CAN WE PROCESS AT ONCE?

Two strings of letters will presented at a time on a computer screen. These two strings may or may not form a word. After a very brief presentation, the two strings of letters will disappear at which point a decision of same or different is required.

By signing below you are agreeing that you have read and understood the Participant Information Sheet and that you agree to take part in this research study.

____________________________ __________________________ ___________

Participant’s Name (printed) Participant’s Signature Date

Laura Wakeford

____________________________________ ________________________________

Printed name of person obtaining consent

Signature of person obtaining consent

Appendix C

Experimental items for Experiment 1. Letter strings containing 4-letters are presented in the first Table, while those containing 6-letters are presented in the second Table. Within these Tables, letter strings are presented as a function of type: high frequency, low frequency, legal nonword and illegal nonword. The word to the left of the back slash represents the central word, while the word to the right represents the peripheral word.

Four Character Words

High Frequency Words	Low Frequency Words	Legal Nonwords	Illegal Nonwords
Army / ermy	womb / worb	lomu / homu	gwdi / iwdi
paid / pait	twin / trin	krin / klin	fsdp / fpdp
body / boly	trek / trik	bron / blon	oajf / oapf
blue / plue	skip / skif	clur / clup	dopm / dopl
data / dita	smog / smeg	eron / fron	daqs / dazs
city / sity	seam / veam	lurn / kurn	hrud / hiud
door / doar	ache / oche	ploc / floc	dsai / psai
poor / poot	ants / auts	smen / smep	pdis / rdis
talk / tark	yawn / rawn	plun / prun	dgnk / dknk
fire / fure	turf / turg	kint / rint	psof / psff
stop / stor	vine / vint	preg / preb	djfj / djfp
open / oren	suck / nuck	blen / bren	dawg / dapg
view / viem	exit / enit	plip / plup	fsdm / fkdm
rest / rist	fade / tade	crun / frun	fsdj / ysdj
unit / unot	germ / gert	uren / tren	hdbf / hrbf
high / hogh	hook / hoak	loup / boup	gtsk / gtsf
show / thow	tart / gart	rult / rula	ptew / puew
dark / derk	maze / waze	baga / biga	twei / twli
like / rike	slap / srap	slon / slun	ptie / ptit
late / lato	spin / spen	iren / irun	gfev / gfpv
wife / bife	wrap / wrat	tibe / tobe	osui / psui
shot / stot	wasp / wamp	plar / plir	pewr / pewg
kind / kild	viva / vila	mena / ment	jkio / jkin
keep / keap	thug / thun	owel / uwel	rskg / rskf
home / hobe	crow / clow	kump / gump	rwom / rkom

word / tord	tidy / tudy	hict / hoct	hedf / hemf
step / sted	stir / spir	twem / swem	fsdh / fgdh
head / heam	sway / swar	drun / druf	yplw / fplw
girl / sirl	bred / bren	plut / Plit	fdfd / rdfd
left / luft	cute / zute	ging / gint	tdsa / tgsa
date / dape	dame / dume	pirt / pirg	traa / trpa
eyes / eyer	frog / frug	fint / finy	trse / trsf
test / pest	vest / pest	wint / wict	esfv / esfg
hour / pour	swan / swam	rapy / ripy	szce / szye
wall / will	curl / cure	himp / hirp	rsfc / rrfc
size / site	peer / peel	slom / slem	utrh / ptrh
mind / mine	lump / limp	quop / quep	asff / fsff
work / worn	quiz / quit	cowe / cobe	gdfg / gsfg
year / near	worm / warm	paye / raye	hugb / huub
walk / talk	sunk / punk	smin / stin	tawv / tawd
food / foot	bump / lump	yarb / yarp	zsdo / zsdd
love / lone	tick / tack	bort / birt	jopf / jogf
hair / pair	bags / bats	grob / gron	jkfo / jrfo
game / tame	bolt / belt	deat / delt	fsoj / gsoj
mean / meat	fake / fate	blit / blim	fsdl / ysdl
east / past	gram / grim	davo / dato	gtjo / gijo
ball / bill	hail / hair	sool / soll	smlp / smrp
role / rule	hunt / hint	rabe / ribe	sfml / sfmb
note / none	knit / knot	vant / vint	cvdm / cvdk
live / life	mice / rice	munt / mont	fvsd / fvod
hope / home	raid / paid	sher / sper	czml / cgml
hall / hill	slug / snug	leva / liva	czxo / pzxo
good / gold	sore / sole	tren / pren	hjfo / fjfo
care / came	wipe / ripe	lork / tork	fcsx / fasx
born / burn	wool / cool	moun / boun	czjo / czho
farm / firm	duck / dusk	marn / marb	gxdl / gxdk
feet / feel	moss / most	rarn / rart	fdas / fdak
five / fine	bust / busy	cand / cald	gdfj / gdvs
deal / peal	poll / pool	ello / elly	fkeo / fjeo
hold / holt	babe / baby	flir / plir	jaop / paop
face / fact	bang / bung	mirp / morp	ghdf / ahdf
lost / list	glue / blue	swon / swin	oxdu / opdu
read / real	lone / line	avil / avit	fsja / fsca
play / pray	arcs / arms	quap / quep	fkop / fkok

Six Character Words

High Frequency Words	Low Frequency Words	Legal Nonwords	Illegal Nonwords
action / actior	vanish / lanish	flippe / fluppe	asjvmo / ksjvmo
attack / altack	wallop / wallip	tarpel / tarper	adsopj / adsopg
closed / clused	danced / dancid	rathew / rathel	dfcjnc / dfsjnc
direct / dirent	anthem / anttem	placky / pracky	sdnklv / sunklv
family / lamily	mortal / mottal	brouse / frouse	wjopfm / wjypfm
ground / grould	coffin / cofkin	wotler / woller	gfsdjd / bfsdjd
health / pealth	deluxe / depuxe	pilner / palner	wokfjo / wokojo
income / incume	freeze / fleeze	quippy / quipsy	rdynjk / rdsnjk
market / marmet	fungal / pungal	blinch / blunch	yssjio / ysdjio
nation / jation	greasy / gleasy	marath / marith	frupop / fruiop
number / nummer	indigo / indiga	punick / penick	fjdkof / ajdkof
pretty / pletty	jumper / jumpor	chaunt / thaunt	fiopza / fiopsa
supply / suptly	kitten / kitsen	richan / rochan	woifml / woyfml
people / prople	litter / littem	sarath / satath	sajdop / lajdop
spirit / sporit	magnet / wagnet	chuman / chiman	jfdfnk / jfdfik
doctor / dactor	oyster / ayster	journt / sournt	sadjpv / sadopv
school / schoon	screws / screwn	oriank / sriank	reopjf / reoujf
public / publin	rocket / ricket	rupart / rumart	fasnkp / faknkp
stress / strest	sponge / spolge	kinkan / kintan	fuphjf / fjphjf
future / fumure	tremor / tremom	smitta / sminta	vfijks / vfbjks
energy / anergy	asylum / psylum	mundle / cundle	fshjio / fshjip
nature / niture	arctic / arttic	prunte / plunte	dskoki / djkoki
person / pesson	escort / escolt	zumbra / lumbra	dsflkn / dsflon
church / chutch	hybrid / hobrid	ticken / tikken	ahiovm / ahiovq
corner / porner	knives / knines	gezzan / gezlan	wjopcv / wjopcp
window / wildow	vexing / mexing	vemant / vemany	gmhuio / gshuio
street / streem	clinic / clipic	polint / wolint	jopasj / jopatj
modern / modean	chords / chorps	smelky / stelky	bcvmrv / bcvmcv
island / islaud	admire / admirp	peason / peaton	djopsm / dropsm
county / coungy	shines / stines	bravem / bravey	czxnyl / czxnml
method / mettod	potter / polter	loopsy / toopsy	fndiom / pndiom
figure / figuro	poison / polson	kitale / kitare	fspnjk / fsanjk
saying / paying	bouncy / bounce	chavin / chaven	dsafio / dsajio
police / polite	willow / wallow	mittle / muttle	uanvfs / uanvfs
latter / tatter	puddle / muddle	pinkle / punkle	fnsaui / fnsqui
things / thinks	warmed / warned	grissy / gristy	iwnpon / iwnjon
simple / sample	worded / worked	funkle / funkly	dfasdj / duasdj
recent / repent	wiring / wiping	ravara / ravata	ysdnko / fsdnko
became / become	wobble / wobbly	vumtin / vumton	assdfs / ausdfs
living / lining	trunks / trumps	smolta / smosta	xdvnjk / xdvnjr
looked / locked	brands / brandy	winkal / wonkal	dasjkl / dasnkl
letter / litter	bolder / colder	quinto / quilto	awjopl / awjopy

forces / forced	chives / chimes	frends / frelds	sajpjo / salpjo
played / placed	crease / grease	poaker / plaker	udjfgu / fdjfgu
strong / strung	divert / divers	rupist / rusist	fsdnio / fsdnip
change / chance	pasted / lasted	kingol / pingol	jdsbnu / fdsbnu
turned / burned	lovers / livers	soupin / sorpin	vjisdo / pjisdo
really / realty	export / extort	goning / goling	spnmkl / sdnmkl
report / resort	cooked / cooled	pethan / pettan	asjiop / lsjiop
walked / walker	detect / deject	poulst / poilst	vxcqmc / vxcnmc
spring / sprang	licked / lacked	kirkel / kirkol	dsajpj / dsajij
father / fatter	rating / raping	nerval / nerpal	waijpl / wuijpl
longer / linger	tasted / wasted	gruppy / grupsy	wtaypd / wtayrd
mother / bother	verses / versus	quoush / quouch	fajiof / fajiot
worked / worker	boiled / bailed	naveer / naveel	wdsnjo / pdsnjo
summer / simmer	Straps / straws	druken / druket	dfniod / dfniom

Appendix D

A representative Participant Information Sheet for experiments involving the

Dr Bouis Eye Tracking Machine (Experiments 2 to 6).

PARTICIPANT INFORMATION SHEET

Parafoveal Processing During Silent Reading

INVITATION TO TAKE PART IN A RESEARCH STUDY

You are being asked to take part in a research study whereby your eye movements will be monitored while reading a collection of short sentences. The data collected will be used for the completion of my PhD, which is supervised by Dr Wayne Murray

PURPOSE OF THE RESEARCH STUDY

The purpose of this research is to look at the way in which people use their eyes while reading. You will be required to read the sentences that appear on the screen in front of you, just as you would normally read a book or paper. On a few occasions, you will be required to answer a comprehension question relating to the sentence you have just read. Participation in this research would benefit myself as it is a necessary component for the completion of my thesis.

TIME COMMITMENT

It is estimated that your participation in the experiment will take approximately one hour. You will only be required to attend one session.

TERMINATION OF PARTICIPATION

Your participation is voluntary. You may decide to stop being a part of the research study at any time without explanation and without penalty. Withdrawing from the study will have no effect for your studies. If you do decide to stop, you will still receive your three course credits. You may also choose to take a break whenever you need one.

RISKS

There are no known risks for you in this study. A sterilised bite bar will be used to control your head movements. During the experiment your eyes will come into contact with an infra-red light in order that we can record your eye movements. Please note that the light levels in the beam are well within health and safety limits.

COST, REIMBURSEMENT AND COMPENSATION

Your participation in this research is voluntary. You will receive three course credits after completion of the testing.

CONFIDENTIALITY/ANONYMITY

The data we collect does not contain any personal information about you. No one will link the data you provided to your identity and name.

FOR FURTHER INFORMATION ABOUT THIS RESEARCH STUDY

Laura Wakeford will be glad to answer your questions about this research at any time. If you want to find out about the final results of this study, you should contact myself, Laura at ljwakeford@dundee.ac.uk.

The University Research Ethics Committee of the University of Dundee has reviewed and approved this research study.

UREC v. 1.9, 15 December 2006

Appendix E

A representative Instruction Sheet for experiments involving the Dr Bouis Eye Tracking Machine (Experiments 2 to 6).

PARTICIPANT INSTRUCTION SHEET

Instructions for Eye Movement Experiment

In this experiment I will monitor your eye movements while you read a number of sentences. Sentences will be presented to you one at a time on a screen. All you need to do is read each sentence normally, as if you were reading it in a book or a newspaper.

Before each sentence appears, a marker (+) will appear briefly on the left hand side of the screen. Please look at this marker as soon as it appears. After a short delay, the sentence will then appear for you to read. It is important that you remember to look at the marker before reading each sentence.

Once you have finished reading each sentence, press the right hand button and you will then be presented with a series of dashes (---). When you are ready for the next sentence, simply press the right hand button again and the next sentence will appear on the screen.

Following a small random sample of sentences you will be asked to answer a simple question that relates directly to the sentence you have just read (this will appear on the screen instead of the dashes). The required response will either be yes or no. If the answer is yes, please press the right hand button, if the answer is no, please press the left hand button. Once you have answered, please fixate on the (+) in preparation for the next sentence.

After every four sentences, I will ask you to look at an array of five numbers displayed on the screen. This is used for calibrating the apparatus and only takes a few seconds but it is important, so please concentrate. Also, throughout the experiment it is vital that you keep as still as possible while the measurements are being made. Please keep your head still and try not to move your arms when you press the response buttons (arrange the buttons so that they lie comfortably under the index finger of each hand).

Before you begin the experiment we will go through some practice items to allow you to become familiar with the task. Please use this time to ask any questions you may have and to make yourself comfortable

Appendix F

A Representative Consent Form for experiments involving the Dr Bouis Eye Tracking Machine (Experiments 2 to 6).

INFORMED CONSENT FORM

PARAFOVEAL PROCESSING DURING SILENT READING

The purpose of this research is to look at the way in which people use their eyes while reading. You will be required to read the sentences that appear on the screen in front of you, just as you would normally read a book or paper. On a few occasions, you will be required to answer a comprehension question relating to the sentence you have just read.

By signing below you are agreeing that you have read and understood the Participant Information Sheet and that you agree to take part in this research study.

____________________________ __________________________ ___________

Participant's Name (printed) Participant's Signature Date

Laura Wakeford

____________________________________ ________________________________

Printed name of person obtaining consent Signature of person obtaining consent

NOTE: The Consent Form should normally be separate from the Participant Information Sheet so that the participant has something they can keep.

UREC v. 1.9, 15 December 2006

Appendix G

Experimental items for Experiment 2. Previews for words n+1 and n+2 are presented in italics and are immediately followed by their associated target words, which are underlined.

Words falling to the left of the forward slash represent the determiner word n+1 previews and targets, while those falling to the right represent the alternative high frequency word n+1 previews and targets. For example:

```
         n      n+1    n+2    n+1    n+2
       taking flm/sia jaknup the/new policy
              det/alt        det/alt
```

The president is taking *flm/sia jaknup* the/new policy from the senate very seriously indeed.

The worried lady called *fbx/imw lwcvnk* the/one friend from the police station after getting arrested.

The brave soldier served *lki/mzv zudwem* the/our nation when he was just eighteen years of age.

The farmer was moving *bfi/hwr riflkz* the/ten cattle to an adjacent field when he saw his neighbour.

The cab driver opened *kfi/lex smakux* the/his window in the taxi because it was hot.

The new teacher had chosen *kbz/skt qicksp* the/old poetry from the 1800s for the students to study.

Yesterday, the kind farmer caught *lfm/hsc kcinxz* the/two horses that had strayed onto his land.

The young girl called *ldi/foi lomsxf* the/her friend on the telephone despite being told not to.

The bad flooding caused *kdo/miz oloimk* the/our church to close for seven weeks.

The midwife showed *hbx/msj xifksn* the/any mother who had just given birth how to handle the baby.

The FBI agent passed *dki/fov azjimf* the/his report to his superior officer to read.

The smart teenager sought *bfo/zsi ifmasn* the/new advice from the Citizens Advice Bureau.

The bride had chosen *dla/viz htiums* the/one flower for her bouquet because she liked the colour.

On request, the receptionist passed *kfu/knu hudbin* the/her letter to her extremely angry boss.

The mean farmer killed *fku/uht zxkfki* the/old cattle himself rather than paying the slaughterhouse.

The clever scientist proved *kfi/feu kfiumj* the/his theory using his newly invented machine.

The local vet helped *flo/irp svrwsh* the/any animal with a sore paw make a full recovery.

The lady had broken *fkv/loz mxvler* the/her window by throwing a brick through it in anger.

The office worker opened *kdm/xvz kihkam* the/one letter addressed to his boss by accident.

The radio presenter played *blx/zop xanzwf* the/any record his listeners had requested.

The new headmaster closed *dfm/frv zntiak* the/his school for two weeks due to storm damage.

Last night, a new neighbour called *klu/nvc huxlui* the/our doctor after her son got sick.

Last night, the tired mother served *bti/kzo lmorsz* the/her dinner and washed up before going to bed.

Luckily, the brave lady pulled *lbs/fvc iafkic* the/her mother from the stream after she had fallen in.

The new fireman called *kfi/uth juhasz* the/all police to the blaze because he thought it could be arson.

The worried farmer walked *dlu/hwa huxktn* the/ten fields in the evening looking for his flock.

The lazy student failed *ldo/loi miqwzl* the/her report on statistics because she did not attend class.

The brave fireman pulled *kba/bue lufuwz* the/two bodies from the house before it collapsed.

In class the teacher failed to answer *fbx/rxm wkafird* the/one student who always asked difficult questions.

The book editor forced *fbo/kwe mochsv* the/his writer to rename the controversial book.

As promised, the pretty girl placed *dbi/fos hukdim* the/her letter on her father's desk.

The hotel employee walked *fku/uht poizhw* the/all guests to their rooms and carried their luggage.

The cafe assistant served *kli/liv nxtkim* the/hot coffee to the lecturer who came in every Thursday.

The young man opened *bdu/tou hifbxc* the/his letter from the court to find out whether he had been fined.

The jockey was taking *fkz/czx lcvaia* the/new horses to the stables when he fell and hurt his leg.

Unfortunately, the young child killed *bdm/kum sowvxk* the/his animal by accidentally overfeeding it.

The office worker had bought *dbw/lsa zwocvt* the/her animal from a place that breeds dogs.

In Canada the famous architect decided to design *flm/mer lcvxzr* the/six houses in a contemporary style.

In Africa, the famine forced *ldo/siq keuokj* the/any family in the remote village to relocate.

The old publisher helped *dbi/dnz mwshuv* the/his writer from the city become extremely famous.

The young hikers walked *lki/vzc imthzp* the/our valley in the evening despite the fading light.

The pretty nurse called *dlo/mcr limhen* the/one doctor to help her medicate the poorly patient.

Yesterday, the young girl helped *fbm/buv sufkic* the/her mother in the kitchen.

Many years ago the vicar had joined *blc/taz ikeuxl* the/his church in the little highland village.

The council employee closed *kfi/svq luofjs* the/any bridge that failed to meet safety standards.

The kind landlord offered *kdo/mas lixrnm* the/new houses to the tenants who had been left homeless.

Last year, the young man shared *fbn/fmn rdkxsu* the/his office with his best friend.

The kind lady was giving *fkw/bzm uhwxni* the/her advice to the boy when he unexpectedly started to cry.

Appendix H

Experimental items for Experiment 3. Previews for words n+1 and n+2 are presented in italics and are immediately followed by their associated target words, which are underlined.

Words falling to the left of the forward slash represent the 4-letter word n+1 previews and targets, while those falling to the right represent 6-letter word n+1 previews and targets. For example:

n	n+1	n+2	n+1	n+2
struck	*hiec/limwkj*	*larmiz*	four/twenty	houses
	4L 6L		4L 6L	

The crashing aeroplane struck *hiec/limwkj larmiz* four/twenty houses in central London after losing power.

The food shortage caused *lupf/foplsa gaxzum* high/higher prices for all food types, including bread and milk.

The talented artist sought *krvc/goilfj cxkiha* five/pretty models to photograph for his art project.

The useless waitress served *uihf/lomsil lursuc* cold/burned dinner to the young couple.

The Swedish girl played *fitk/osmhux aguilz* ball/winter sports after school with her friends.

The boy refused to follow *kanv/xanghr inlumn* firm/simple orders from his parents to stay in his room.

Unfortunately, the virus killed *xamu/lfanlg girqhi* nine/thirty people on the plane from India to London.

The new judge picked *lomc/locvhg kvhuxv* five/twenty ladies to proceed to the next stage of the beauty contest.

The patient's operation had caused *guis/lihkam cmauer* poor/better vision in her right eye only.

In Iraq the army doctor taught *uarl/lkanfg lauigz* sick/thirty troops how to reapply their own bandages.

The exploding bomb killed *ierg/luihnm bwewqi* many/twelve troops who had been guarding the American Embassy.

The young builder picked *xcvw/wmlumw larfhc* some/entire fields to develop with the money he had inherited.

The famous orchestra played *somi/lrcilg qanxvm* nine/twenty pieces to the Queen during her visit.

The old psychologist forced *uxol/licvnm grijhw* sick/famous people to re-evaluate their priorities.

The tired teacher marked *uizg/uxfimd juqizx* many/school papers before deciding to give up for the night.

The new charity sought *zxan/ehvmep hukzxa* nice/strong ladies to help other women overcome domestic abuse.

The lady had bought *wsnx/xionvz ghimdw* more/summer plants for her new flat to make it feel more homely.

After class, the music teacher picked *lisn/hcxtmz jewcqv* four/twelve groups to compete in the battle of the bands.

The journal editor had chosen *usxj/lfanbq gnyniz* many/thirty papers to publish before he decided to resign.

Yesterday, the estate agent showed *hasm/ivbmva faiznv* four/modern houses to the family who were relocating.

The old judge forced *asnl/ihramc girqki* rich/eleven people to pay damages to the poor victims.

Yesterday, the old man walked *ewuk/jnihkg pixnhr* rich/pretty guests through the grounds of the castle.

After his crash, the teenager waited *ioej/ifcmaw xzckfi* many/eleven months before he was allowed to drive again.

In the lab, the new psychologist showed *qimn/filham wnvsip* poor/better memory in a number of circumstances.

The gym instructor showed *luma/irkcoz giajka* five/active people how to work the machines in his gym.

Yesterday, the landscape gardener picked *oanz/pudhim jkvxko* nice/yellow plants to border the pond and driveway.

In the hotel, the armed robber forced *camv/flamkp paizdv* nine/thirty guests to hand over all their money.

The kind nurse helped *remh/birkna kubamn* sick/twelve ladies go home for Christmas by offering home visits.

The careless boy killed *lamg/krhfha qheidm* tiny/little plants in the garden by flattening them with his bike.

At the board meeting, the man raised *lanr/wamuxf vxnoar* five/recent issues he had encountered with the employee.

The kind man helped *vazi/ilavmj limwic* some/strong horses find new homes after his farm went bankrupt.

At the army training centre, the captain pushed *nazc/huakmi hameqn* some/twelve troops into making a complaint.

In France, the old employer bought *larg/hoewhg kiazcx* tiny/family houses for his workers to live in.

The young traveller had chosen *wani/dirnzr xndamr* some/famous cities to visit during his travels to the states.

The tour guide showed *baoc/janklg gkinvm* five/pretty places within the brochure the tourists could visit.

Unfortunately, the camp site forgot to charge *texn/iharwm jiandz* four/eleven guests for the gas they had used.

The female detective had to search *saxv/ndirmn fukauz* nine/eleven ladies after the money went missing.

The pretty lady raised *hain/dimzdj banziv* four/twenty horses from a very young age without any financial help.

The boss wanted to supply *irnl/zlamij irdlan* weak/strong coffee to his employees to maintain motivation.

In America, the recession had struck *gaun/himyin nuhaxn* poor/larger cities the hardest according to the report.

The young child caught *esrg/huflka fonizc* many/little leaves as they fell from the old oak tree.

The landlady decided to charge *lekl/adcuri pranlx* both/eleven guests for the damage they had caused.

The winter had caused *biaj/damjiz crpkbz* long/longer nights because the daylight started fading at 4pm.

The self-help book taught *anvl/hamikg jaigha* rich/lonely people how to avoid being taken advantage of.

The fashion agency taught *hubd/ievcuk urbikz* tall/normal models how to be successful on the catwalks.

The grumpy policeman decided to charge *ruom/handim hxkcom* nine/twelve ladies with being drunk and disorderly.

For the filming, the director had chosen *lipr/danpld kapbfa* huge bright lights to light up the lake.

The financial crisis caused *fapi/ahopkb kcimnv* huge/slight losses in jobs over the course of the year.

Appendix I

A representative example of a Participant Information Sheet used for plausibility ratings.

PARTICIPANT INFORMATION SHEET

PLAUSIBILITY EFFECTS ON PARAFOVEAL PREVIEW

INVITATION TO TAKE PART IN A RESEARCH STUDY

You are being asked to take part in a research study in which you will be asked to rate a series of sentence fragments according to how plausible or grammatical you find them. The data collected will be used for the completion of my PhD, which is supervised by Dr Wayne Murray.

PURPOSE OF THE RESEARCH STUDY

The purpose of this research is to provide some predictability ratings on some item sets that will be used for an eye tracking study. Your task is to read each of the fragments and give them a rating on a seven point scale of plausibility. Alternatively, if you feel the sentence is completely unnatural or ungrammatical, you simply have to circle 'U' and do not have to provide a plausibility rating. Participation in this research would benefit myself as it is a necessary component for the completion of my thesis.

TIME COMMITMENT

It is estimated that your participation in the experiment will take approximately one hour. You will only be required to attend one session.

TERMINATION OF PARTICIPATION

Your participation is voluntary. You may decide to stop being a part of the research study at any time without explanation and without penalty. Withdrawing from the study will have no effect for your studies. If you do decide to stop, you will still receive your three course credits. You may also choose to take a break whenever you need one.

RISKS

There are no known risks for you in this study.

COST, REIMBURSEMENT AND COMPENSATION

Your participation in this research is voluntary. You will receive three course credits after completion of the testing.

CONFIDENTIALITY/ANONYMITY
The data we collect do not contain any personal information about you. No one will link the data you provided to your identity and name.

FOR FURTHER INFORMATION ABOUT THIS RESEARCH STUDY
Laura Wakeford will be glad to answer your questions about this research at any time. If you want to find out about the final results of this study, you should contact myself, Laura at ljwakeford@dundee.ac.uk.

The University Research Ethics Committee of the University of Dundee has reviewed and approved this research study.

UREC v. 1.9, 15 December 2006

Appendix J

A representative example of an Instruction Sheet for participants completing

the plausibility ratings followed by an example item.

Instructions for Ratings of “Plausibility” and “Naturalness”

I want you to read each of the following sentence fragments and give them a rating for “plausibility”. By this I mean the ordinariness or likelihood of the event being described by the fragment actually happening or being true. Thus a plausible fragment will describe a very ordinary event, which has a high probability of occurring in everyday life, whereas an implausible fragment will describe a very bizarre or unexpected event, which is not very likely to occur. For example, the fragment:

The bird flew…

is very plausible, whereas the fragment…

The hedgehog asked…

is very implausible.

Your task is to read each of the sentence fragments and give them a rating on 7-point scale of plausibility on which a score of 7 corresponds to a highly plausible fragment, while a score of 1 should be given to a highly implausible fragment, with other scores representing gradings on the scale at equal intervals in between.

If you read the sentence fragment and decide that it is either ungrammatical or completely unnatural, please DO NOT give the sentence a plausibility rating; instead circle the ‘U’, which indicates you believe it to be either ungrammatical or completely unnatural.

Thus, for each sentence fragment you should either circle a number corresponding to your estimate of its plausibility or circle the ‘U’ if you believe it is ungrammatical or completely unnatural. Choose just one of the numbers or the ‘U’ on the same line as the item.

<u>Examples:</u>

***"The happy student passed the four…"* is pretty plausible,
so you would circle a high number on the scale.**

***"The happy student passed the dust…"*
is rather less likely, so you would circle a low number.**

***"The happy student passed the receive…"*
would probably be given a 'U'.**

Please treat each fragment as the beginning of a sentence, and rate the plausibility of just this fragment (ignoring anything that might follow it and affect how likely or otherwise it may seem).

It is important you do not think too long before circling a rating. Just quickly read the item and give your first impression.

Example Item:

1 2 3 4 5 6 7 U   The security guard wanted advice …

Appendix K

A representative example of a Consent Form used for the plausibility ratings.

INFORMED CONSENT FORM

PLAUSIBILITY EFFECTS ON PARAFOVEAL PREVIEW

The purpose of this research is to look at the way in which people use their eyes while reading, for which you will be providing the plausibility ratings. Your task is to read each of the fragments and give a rating on a seven point scale of plausibility on which a score of 7 corresponds to a highly plausible fragment, while a score of 1 should be given to a highly implausible fragment, with the other scores representing gradings on the scale at equal intervals in between. Alternatively, if you feel the sentence is completely unnatural or ungrammatical, you simply have to circle 'U' and do not have to provide a plausibility rating.

By signing below you are agreeing that you have read and understood the Participant Information Sheet and that you agree to take part in this research study.

____________________________ __________________________ ___________

Participant's Name (printed) Participant's Signature Date

Laura Wakeford

____________________________________ ________________________________

Printed name of person obtaining consent Signature of person obtaining consent

NOTE: The Consent Form should normally be separate from the Participant Information Sheet so that the participant has something they can keep.

UREC v. 1.9, 15 December 2006

Appendix L

A representative example of the Participant Information Sheet used for predictability ratings.

PARTICIPANT INFORMATION SHEET

PREDICTABILITY RATINGS

INVITATION TO TAKE PART IN A RESEARCH STUDY

You are being asked to take part in a research study to provide predictability norms for an eye tracking experiment. The data collected will contribute towards my Ph.D. research, which is being supervised by Dr Wayne Murray.

PURPOSE OF THE RESEARCH STUDY

The purpose of this research is to provide predictability norms for some experimental items used in an eye tracking experiment. You will be required to read a series of sentence fragments; these fragments will always constitute the beginnings of sentences for which you have to predict what the next word will be. This research will benefit me as it is a necessary component of my Ph.D.

TIME COMMITMENT

This experiment will take half an hour to complete and will take place in the Annex Lab.

COST, REIMBURSEMENT AND COMPENSATION

Your participation in this study is voluntary. You will receive 2 course credits after completion of the testing.

RISKS

There are no known risks for you in this study.

TERMINATION OF PARTICIPATION

You may decide to stop being a part of the research study at any time without explanation and without penalty.

CONFIDENTIALITY/ANONYMITY

The data collected do not contain any personal information about you. No one will be able to link the data you provided to your identity and name. The data will be seen only by the researchers and will not be made available to anyone else. If this research is published, you will not be identifiable in any way.

FOR FURTHER INFORMATION ABOUT THIS RESEARCH STUDY

I will be glad to answer your questions about this study at any time. You may contact me at ljwakeford@dundee.ac.uk or via my School phone number: (01382) 388129. If you want to find out about the final results of this study, you should contact me via email.

The University Research Ethics Committee of the University of Dundee has reviewed and approved this research study.

Appendix M

A representative example of an Instruction Sheet for participants completing the predictability ratings followed by an example item.

PREDICTABILITY RATINGS

The following pages contain a series of sentence fragments. Your task is to continue each sentence with the first word that comes to mind. To be clear, you do not need to complete the sentence; you simply need to write down what you think the next word will be. Please do not spend too much time on each item; there are no "correct" answers, we just want to know THE FIRST WORD THAT COMES INTO YOUR MIND.

For Example:

Tim had been playing *football*

Or

The postman had collected *the*

Or

Rebecca was taking care to *carry*

Example Item:

The exploding bomb had killed ______________________

Appendix N

A representative example of a Consent Form used for the predictability ratings.

CONSENT FORM

Predictability Ratings

You will be required to read a series of sentence fragments; these fragments will always constitute the beginnings of sentences for which you have to predict what the next word will be.

By signing below you are indicating that you have read and understood the Participant Information Sheet and that you agree to take part in this research study.

__________________________________ __________________

Participant's signature Date

____

Participant's name

__________________________________ __________________

Signature of person obtaining consent Date

Laura Wakeford

____

Name of person obtaining consent

Appendix O

Experimental items for Experiment 4. The previews for word n+1 are presented in italics and are immediately followed by their associated target word, which is underlined.

The italicised words represent the plausible, anomalous and illegal nonword previews, respectively. For example:

n n+1 n+1

making *coffee/caught/fumeio* dinner

Plaus / Anom / Illegal

The mother was making *coffee/caught/fumeio* dinner in the kitchen for her two children and her husband.

Many years ago, the man helped *slaves/inches/jmnpki* people in the concentration camps to escape.

The adverse conditions did not affect *ladies/forget/uvhwen* cities in the very south of the country.

At the festival, the man had chosen *groups/struck/nzcala* events that he had really wanted to watch.

The property developer bought *fields/stress/tiamnx* houses in the countryside to renovate into flats.

The young receptionist showed *papers/region/pcnalw* guests to their rooms in the exclusive hotel.

The art dealer raised *issues/create/giemzn* prices on the pair of antique paintings after much persuasion.

The hotel owner had to charge *people/breath/psinle* guests for the damage caused during their short stay.

The medieval town looked *little/market/juxlhg* pretty from the top of the hill in the Somerset countryside.

The contagious virus killed *troops/regard/qamgha* people on the aeroplane from London to Istanbul.

The poor orphan wanted *advice/listen/liwahp* family to come home to rather than an empty lonely house.

The man's bad behaviour had caused *ladies/bought/gnapkw* people to make official complaints immediately.

The tropical storm caused *clouds/coming/wnfoah* relief after a long period of draught on the island.

The ruthless criminal forced *horses/before/yimgki* people into the cellar as he raided their property.

The financial advisor had been giving *papers/become/otsmna* advice to the young couple about mortgages.

The bar tender worked *events/letter/oapkfz* nights at the pub before quitting due to the antisocial hours.

The security guard wanted *coffee/strike/olmwun* advice on his employment rights after he was attacked.

The exploding bomb had killed *people/circle/lunmge* troops who had been guarding the American Embassy.

The dying man had wanted *advice/cannot/imgpiz* oxygen to help him breathe during his final few hours.

The artist's house looked *pretty/glance/iwkeom* modern compared to all his neighbour's period properties.

The sleepy teacher marked *poetry/motion/giqime* papers in her office before deciding to go to her bed.

The bright student shared *poetry/itself/uohbiw* coffee with her best friend while they were revising.

The indoor tree needed *leaves/number/foqblu* lights to go with all the tinsel and pretty baubles.

The body builder wanted *dinner/assume/ieovbm* muscle so that he could look like the men in the magazines.

The detectives needed to search *guests/survey/gkweno* places in the local town for the missing drugs.

The unimpressed punter was served *dinner/window/uxbkum* coffee in the dirty and dingy restaurant.

The dedicated estate agent had showed *ladies/matter/kwimsn* houses to the incredibly fussy house buyers.

The crashing aeroplane struck *people/theory/tanucn* houses as it lost power and plummeted to the ground.

The hospital porter was taking *plants/should/fekmni* bodies to the morgue when he slipped on some water.

The hard working man had waited *tables/camera/unikta* <u>months</u> to get his allotment in the small village.

The brave fireman pulled *guests/debate/kuhzni* <u>ladies</u> from the burning house just before it collapsed.

The talented photographer showed *guests/minute/uiopnm* <u>images</u> to the paying client, who loved them.

The rich man had bought *houses/burden/swhbkm* <u>cattle</u> in the rural area to start up his own dairy farm.

The excited bride had chosen *plants/please/geikep* <u>poetry</u> to be read at her forthcoming wedding ceremony.

The grumpy policeman decided to charge *groups/reveal/jsvgha* <u>people</u> in the pub with antisocial behaviour.

The new diet had helped *people/tongue/hukoui* <u>ladies</u> to lose lots of weight in just one month.

The new farmer filled *houses/thrown/krnkhc* <u>fields</u> on his land with the best corn he could buy.

The fashion agency had needed *ladies/though/xcbufa* <u>models</u> for the catwalk in Paris during fashion week.

The gym instructor had showed *people/reader/ykweim* <u>places</u> to his boss that posed a health and safety risk.

The footballer had become *strong/people/jnesma* <u>famous</u> by the time he had reached his early twenties.

The confused tenant needed *dinner/begins/itmsna* <u>advice</u> from the Citizen's Advice Bureau about his rights.

The young boy gained *height/person/xmuplf* <u>weight</u> in the summer holidays because he spent it watching TV.

The rich aristocrat bought *tables/unable/lneuma* <u>horses</u> from the famous breeder for his favourite jockey.

The new gun laws helped *people/column/oahemi* <u>cities</u> in the south of the country to reduce gun crime.

The old athlete played *sports/editor/neirka* <u>events</u> in the tournament despite his age and still won them.

The successful farmer had needed *horses/stared/kzmkfe* <u>fields</u> in which to keep his free range chickens.

The tired waiter showed *groups/oxygen/fikiux* <u>ladies</u> to their tables in the extremely busy restaurant.

The journal editor picked *errors/wonder/qugewm* <u>papers</u> that he wished to publish in the famous journal.

The clever psychologist showed *methods/passage/otofewn* studies to his students that were fatally flawed.

When in town, the girl picked *clothes/disease/lneimlv* friends that she wanted to attend her birthday party.

The charity for the homeless needed *support/whether/ohiklei* clothes to give to those living on the streets.

The reckless boy had broken *windows/message/humjuna* fingers after he fell from his bike while doing a stunt.

The increased opening hours caused *workers/happens/kienmvc* tension in the office between the managers.

The credit crunch had struck *parents/bedroom/wikuikz* schools that were council funded in central London.

The battling army needed *support/fingers/oaigimw* weapons if they stood a chance of winning the terrible war.

The couple had chosen *friends/started/xvksnlz* schools in the area they would consider sending their son to.

The generous bursary had helped *schools/version/jcwmnle* parents in the local area to buy essential books.

The girl was teased for having *answers/becomes/qeisnkm* parents that were unable to pay for a summer holiday.

The elderly ladies shared *clothes/contain/ckernmw* stories of their childhood with one another all evening.

The mother was taking *letters/brought/lnsmoka* friends to the restaurant when her car ran out of petrol.

The hospital DJ played *records/clothes/cluvnma* stories on the radio that some of his listeners had written.

The new charity needed *members/appears/nxmlwui* workers to be on call for emergencies during the night shift.

The business man sought *records/suppose/kajumai* figures from his bookkeeper before going to see the tax man.

The young couple became *parents/storage/hxcmxlw* friends after going on holiday together to Switzerland.

The football team wanted *success/element/iwqpinh* support from their manager rather than his usual bullying.

The author hoped her sons would become *friends/purpose/msokimo* writers when they were older just like her.

The poor employee was giving *answers/believe/nzmxeio* reasons as to why he was late when he was sacked.

The power outage did not affect *schools/realise/xkowmlo* streets in the north of the city during the storm.

The social worker listed *factors/between/neczxmn* reasons as to why the child should go into foster care.

The school inspector closed *windows/animals/mckuekm* schools in the area that he argued were under performing.

The old receptionist opened *windows/persons/hskhoao* letters that were addressed to her manager every day.

The oil tycoon was having *trouble/require/nzmxcie* success in the south of America after striking oil twice.

The kind woman helped *animals/silence/hwemnla* friends in the village to learn the art of oil painting.

The smart policeman sought *records/imagine/urmnsnx* reasons as to why he should not prosecute the suspect.

The black plague struck *workers/getting/olunoki* streets on the outskirts of town where the peasants lived.

The PE teacher taught *methods/windows/mksnxcv* classes in the local village school where his wife also taught.

The newspaper had to report *stories/operate/rwmskli* results on the local elections the very next day.

The teacher was giving *flowers/tension/vfusmna* classes on self defence to help the vulnerable community.

The men were giving *flowers/operate/oiugmni* weapons to the drug runners who were also ruthless criminals.

The lady had to return *animals/nuclear/vheltim* clothes to the shops because they were far too small for her.

The dyslexic child was having *success/include/cnqpiur* support with his writing skills in his new school.

The Olympian was having *success/instead/kcmxhti* trouble trying to perfect his high jump after his injury.

The supply teacher taught *classes/evident/tcmkiaj* history in the past but she hated it as it bored her.

The teenager wanted *clothes/measure/hnmclia* freedom from his parents so he moved out and got himself a job.

The lively child had wanted *animals/husband/xlianms* stories to be read to her before going to sleep.

The sleeping cat looked *strange/attempt/uikltih* settled as she slept on the rug by the roaring open fire.

The head teacher wanted *history/follows/cmnzoia* science to be taught at a younger age in her school.

The helicopter pilot was flying *workers/failure/rmajiom* weapons to the stranded soldiers as ordered.

The old women picked *clothes/receive/htmonie* flowers from the beautiful gardens of the grand manor house.

The athletic man had broken *flowers/explain/cxzoila* records at the Olympic Games despite being ill.

The caretaker of the building locked *offices/balance/iomlnmu* windows in the factory while he was on duty.

The journalist was asked to report *tragedy/provide/mfuourw* stories on the devastating conflicts abroad.

The sisters had become *friends/prevent/euhwila* artists in their adult years, just like their father had.

The unusually strong winds had broken *flowers/thought/smnlezm* windows in the old derelict farmhouse.

The freezing weather killed *animals/herself/kfunmia* flowers in the lady's garden much to her disappointment.

The new company tried to obtain *support/achieve/iklnxmz* offices in the prestigious building in London.

Appendix P

Experimental items for Experiment 5. The previews for words n+1 and n+2 are presented in italics and are immediately followed by their associated target words, which are underlined.

The first italicised word is the nonword n+1 preview, while the following three italicised words, which are separated by forward slashes, are the plausible, anomalous and illegal nonword n+2 previews, respectively. For example:

n n+1 n+2 n+1 n+2

showed *lwfk papers/region/pcnalw* both guests

Plaus / Anom / Illegal

The young receptionist showed *lwfk papers/region/pcnalw* both guests to their rooms in the exclusive hotel.

The exploding bomb had killed *uerp people/circle/lunmge* many troops who had been guarding the American Embassy.

The sleepy teacher marked *xcmo poetry/motion/giqime* some papers in her office before deciding to go to her bed.

The mother was taking *bwrx letters/brought/lnsmoka* four friends to the restaurant when her car ran out of petrol.

The unimpressed punter was served *nwkl dinner/window/uxbkum* cold coffee in the dirty and dingy restaurant.

The credit crunch had struck *giem parents/bedroom/wikuikz* poor schools that were council funded in central London.

The new diet had helped *zcvl people/tongue/hukoui* sick ladies to lose lots of weight in just one month.

The freezing weather killed *irkl animals/herself/kfunmia* wild flowers in the lady's garden much to her disappointment.

The football team wanted *nwek success/element/iwqpinh* real support from their manager rather than his usual bullying.

The crashing aeroplane struck *vmig people/theory/tanucn* many houses as it lost power and plummeted to the ground.

The property developer bought *loqw fields/stress/tiamnx* huge houses in the countryside to renovate into flats.

The artist's house looked *rnug pretty/glance/iwkeom* very modern compared to all his neighbour's period properties.

The poor orphan wanted *mzvx advice/listen/liwahp* some family to come home to rather than an empty lonely house.

The lady had to return *nmio animals/nuclear/vheltim* some clothes to the shops because they were far too small for her.

The PE teacher taught *nscp methods/windows/mksnxcv* many classes in the local village school where his wife also taught.

The tropical storm caused *tcpa clouds/coming/wnfoah* huge relief after a long period of draught on the island.

The social worker listed *hnma factors/between/neczxmn* five reasons as to why the child should go into foster care.

The girl was teased for having *gnuo answers/becomes/qeisnkm* poor parents that were unable to pay for a summer holiday.

The old athlete played *hnld sports/editor/neirka* both events in the tournament despite his age and still won them.

The teacher was giving *koae flowers/tension/vfusmna* free classes on self defence to help the vulnerable community.

The caretaker of the building locked *lukl offices/balance/iomlnmu* both windows in the factory while he was on duty.

The man's bad behaviour had caused *mruf ladies/bought/gnapkw* rich people to make official complaints immediately.

The helicopter pilot was flying *mxrw workers/failure/rmajiom* some weapons to the stranded soldiers as ordered.

The body builder wanted *izio dinner/assume/ieovbm* more muscle so that he could look like the men in the magazines.

The new charity needed *jxvk members/appears/nxmlwui* good workers to be on call for emergencies during the night shift.

At the festival, the man had chosen *tsvn groups/struck/nzcala* live events that he had really wanted to watch.

The detectives needed to search *vmcn guests/survey/gkweno* nine places in the local town for the missing drugs.

The young couple became *prcl parents/storage/hxcmxlw* good friends after going on holiday together to Switzerland.

The new farmer filled *dnkl houses/thrown/krnkhc* <u>both fields</u> on his land with the best corn he could buy.

The new gun laws helped *qunm people/column/oahemi* <u>poor cities</u> in the south of the country to reduce gun crime.

The indoor tree needed *smno leaves/number/foqblu* <u>more lights</u> to go with all the tinsel and pretty baubles.

The hospital DJ played *ynul records/clothes/cluvnma* <u>good stories</u> on the radio that some of his listeners had written.

The fashion agency had needed *krtf ladies/though/xcbufa* <u>tall models</u> for the catwalk in Paris during fashion week.

The newspaper had to report *trkd stories/operate/rwmskli* <u>full results</u> on the local elections the very next day.

The couple had chosen *pcvl friends/started/xvksnlz* <u>good schools</u> in the area they would consider sending their son to.

The confused tenant needed *ncvi dinner/begins/itmsna* <u>more advice</u> from the Citizen's Advice Bureau about his rights.

The school inspector closed *urmg windows/animals/mckuekm* <u>many schools</u> in the area that he argued were under performing.

The sisters had become *hcnv friends/prevent/euhwila* <u>fine artists</u> in their adult years, just like their father had.

The old receptionist opened *uenp windows/persons/hskhoao* <u>many letters</u> that were addressed to her manager every day.

The dying man had wanted *uznx advice/cannot/imgpiz* <u>more oxygen</u> to help him breathe during his final few hours.

The excited bride had chosen *vnom plants/please/geikep* <u>some poetry</u> to be read at her forthcoming wedding ceremony.

The journal editor picked *dnuo errors/wonder/qugewm* <u>five papers</u> that he wished to publish in the famous journal.

The smart policeman sought *nrwz records/imagine/urmnsnx* <u>more reasons</u> as to why he should not prosecute the suspect.

The power outage did not affect *nizp schools/realise/xkowmlo* <u>many streets</u> in the north of the city during the storm.

The mother was making *zmuv coffee/caught/fumeio* <u>some dinner</u> in the kitchen for her two children and her husband.

The dyslexic child was having *fuon success/include/cnqpiur* <u>less support</u> with his writing skills in his new school.

The sleeping cat looked *mrup strange/attempt/uikltih* <u>very settled</u> as she slept on the rug by the roaring open fire.

The new company tried to obtain *zrei support/achieve/iklnxmz* more offices in the prestigious building in London.

The increased opening hours caused *nrvu workers/happens/kienmvc* some tension in the office between the managers.

The young boy gained *mrwn height/person/xmuplf* some weight in the summer holidays because he spent it watching TV.

The grumpy policeman decided to charge *cmek groups/reveal/jsvgha* most people in the pub with antisocial behaviour.

The medieval town looked *mneg little/market/juxlhg* very pretty from the top of the hill in the Somerset countryside.

The Olympian was having *omnl success/instead/kcmxhti* such trouble trying to perfect his high jump after his injury.

The gym instructor had showed *tnrw people/reader/ykweim* five places to his boss that posed a health and safety risk.

The rich aristocrat bought *lnlf tables/unable/lneuma* both horses from the famous breeder for his favourite jockey.

The security guard wanted *vxzv coffee/strike/olmwun* some advice on his employment rights after he was attacked.

The lively child had wanted *ncup animals/husband/xlianms* many stories to be read to her before going to sleep.

The bright student shared *kmrz poetry/itself/uohbiw* fine coffee with her best friend while they were revising.

The rich man had bought *xnvz houses/burden/swhbkm* nine cattle in the rural area to start up his own dairy farm.

The hospital porter was taking *tncd plants/should/fekmni* dead bodies to the morgue when he slipped on some water.

The oil tycoon was having *ncuk trouble/require/nzmxcie* much success in the south of America after striking oil twice.

The old women picked *juek clothes/receive/htmonie* pink flowers from the beautiful gardens of the grand manor house.

The men were giving *mnrk flowers/operate/oiugmni* real weapons to the drug runners who were also ruthless criminals.

The teenager wanted *cvzx clothes/measure/hnmclia* more freedom from his parents so he moved out and got himself a job.

The elderly ladies shared *ouaz clothes/contain/ckernmw* nice stories of their childhood with one another all evening.

The hotel owner had to charge *kcfl people/breath/psinle* both guests for the damage caused during their short stay.

The contagious virus killed *oiuf troops/regard/qamgha* weak people on the aeroplane from London to Istanbul.

The financial advisor had been giving *brmi papers/become/otsmna* free advice to the young couple about mortgages.

The tired waiter showed *cmvl groups/oxygen/fikiux* rich ladies to their tables in the extremely busy restaurant.

The battling army needed *vnio support/fingers/oaigimw* more weapons if they stood a chance of winning the terrible war.

The clever psychologist showed *nrez methods/passage/otofewn* some studies to his students that were fatally flawed.

The talented photographer showed *kibl guests/minute/uiopnm* both images to the paying client, who loved them.

The black plague struck *mcnu workers/getting/olunoki* nine streets on the outskirts of town where the peasants lived.

The generous bursary had helped *rcnp schools/version/jcwmnle* many parents in the local area to buy essential books.

The business man sought *nsvp records/suppose/kajumai* many figures from his bookkeeper before going to see the tax man.

The charity for the homeless needed *eruo support/whether/ohiklei* more clothes to give to those living on the streets.

The head teacher wanted *cuiv history/follows/cmnzoia* more science to be taught at a younger age in her school.

The adverse conditions did not affect *hmnw ladies/forget/uvhwen* four cities in the very south of the country.

The art dealer raised *kzlf issues/create/giemzn* both prices on the pair of antique paintings after much persuasion.

The dedicated estate agent had showed *krnz ladies/matter/kwimsn* four houses to the incredibly fussy house buyers.

The kind woman helped *kcrw animals/silence/hwemnla* five friends in the village to learn the art of oil painting.

The supply teacher taught *cnvm classes/evident/tcmkiaj* some history in the past but she hated it as it bored her.

The bar tender worked *hrvp events/letter/oapkfz* busy nights at the pub before quitting due to the antisocial hours.

The author hoped her sons would become *kuox friends/purpose/msokimo* fine writers when they were older just like her.

The footballer had become *mzcg strong/people/jnesma* very famous by the time he had reached his early twenties.

The unusually strong winds had broken *kukl flowers/thought/smnlezm* both windows in the old derelict farmhouse.

Many years ago, the man helped *zxcv slaves/inches/jmnpki* some people in the concentration camps to escape.

The successful farmer had needed *kiqw horses/stared/kzmkfe* huge fields in which to keep his free range chickens.

The ruthless criminal forced *mzcs horses/before/yimgki* nine people into the cellar as he raided their property.

The athletic man had broken *naej flowers/explain/cxzoila* many records at the Olympic Games despite being ill.

The hard working man had waited *ncmp tables/camera/unikta* many months to get his allotment in the small village.

The reckless boy had broken *kcun windows/message/humjuna* four fingers after he fell from his bike while doing a stunt.

When in town, the girl picked *znui clothes/disease/lneimlv* more friends that she wanted to attend her birthday party.

The journalist was asked to report *rwne tragedy/provide/mfuourw* more stories on the devastating conflicts abroad.

The brave fireman pulled *eomg guests/debate/kuhzni* many ladies from the burning house just before it collapsed.

The poor employee was giving *ymal answers/believe/nzmxeio* good reasons as to why he was late when he was sacked.

Appendix Q

Experimental items for Experiment 6. The previews for word n+2 are presented in italics and are immediately followed by their associated target words, which are underlined.

The italicised words are the plausible, anomalous and illegal nonword previews, respectively. For example:

n n+1 n+2 n+2

showed the *papers/beside/pcnalw* guests

Plaus / Anom / Illegal

Many years ago, the man helped six *slaves/inches/jmnpki* people in the concentration camps to escape.

The adverse conditions did not affect any *ladies/forget/uvhwen* cities in the very south of the country.

At the festival, the man had chosen six *groups/struck/nzcala* events that he had really wanted to watch.

The property developer bought the *fields/stress/tiamnx* houses in the countryside to renovate into flats.

The young receptionist showed the *papers/beside/pcnalw* guests to their rooms in the exclusive hotel.

The art dealer raised the *issues/create/giemzn* prices on the pair of antique paintings after much persuasion.

The security guard wanted new *coffee/nodded/olmwun* advice on his employment rights after he was attacked.

The hotel owner had to charge his *people/breath/psinle* guests for the damage caused during their short stay.

The medieval town looked too *little/walked/juxlhg* pretty for the army to demolish in the Cornish countryside.

The contagious virus killed the *troops/regard/qamgha* people on the aeroplane from London to Istanbul.

The poor orphan wanted her *advice/listen/liwahp* family to come home to rather than an empty lonely house.

The man's bad behaviour had caused fat *ladies/bought/gnapkw* people to make official complaints immediately.

The tropical storm caused big *clouds/coming/wnfoah* relief after a long period of draught on the island.

The ruthless criminal forced the *horses/before/yimgki* people into the cellar as he raided their property.

The financial advisor had been giving the *papers/become/otsmna* advice to the young couple about mortgages.

The bar tender worked ten *events/letter/oapkfz* nights at the pub before quitting due to the antisocial hours.

The exploding bomb had killed the *people/circle/lunmge* troops who had been guarding the American Embassy.

The dying man had wanted the *advice/cannot/imgpiz* oxygen to help him breathe during his final few hours.

The artist's house looked too *pretty/glance/iwkeom* modern compared to all his neighbour's period properties.

The mother was making the *coffee/caught/fumeio* dinner in the kitchen for her two children and her husband.

The sleepy teacher marked the *poetry/motion/giqime* papers in her office before deciding to go to her bed.

The bright student shared her *poetry/itself/uohbiw* coffee with her best friend while they were revising.

The indoor tree needed the *leaves/toward/foqblu* lights to go with all the tinsel and pretty baubles.

The body builder wanted his *dinner/assume/ieovbm* muscle so that he could look like the men in the magazines.

The detectives needed to search six *guests/thanks/gkweno* places in the local town for the missing drugs.

The unimpressed punter was served one *dinner/window/uxbkum* coffee in the dirty and dingy restaurant.

The dedicated estate agent had showed the *ladies/seemed/kwimsn* houses to the incredibly fussy house buyers.

The crashing aeroplane struck the *people/theory/tanucn* houses as it lost power and plummeted to the ground.

The hospital porter was taking the *plants/should/fekmni* bodies to the morgue when he slipped on some water.

The hard working man had waited ten *tables/camera/unikta* <u>months</u> to get his allotment in the small village.

The brave fireman pulled the *guests/debate/kuhzni* <u>ladies</u> from the burning house just before it collapsed.

The talented photographer showed her *guests/minute/uiopnm* <u>images</u> to the paying client, who loved them.

The rich man had bought ten *houses/burden/swhbkm* <u>cattle</u> in the rural area to start up his own dairy farm.

The excited bride had chosen the *plants/please/geikep* <u>poetry</u> to be read at her forthcoming wedding ceremony.

The grumpy policeman decided to charge the *groups/reveal/jsvgha* <u>people</u> in the pub with antisocial behaviour.

The new diet had helped fat *people/tongue/hukoui* <u>ladies</u> to lose lots of weight in just one month.

The new farmer filled all *houses/thrown/krnkhc* <u>fields</u> on his land with the best corn he could buy.

The fashion agency had needed the *ladies/though/xcbufa* <u>models</u> for the catwalk in Paris during fashion week.

The gym instructor had showed two *people/reader/ykweim* <u>places</u> to his boss that posed a health and safety risk.

The footballer had become too *strong/states/jnesma* <u>famous</u> to be living in the big city without bodyguards.

The confused tenant needed the *dinner/begins/itmsna* <u>advice</u> from the Citizen's Advice Bureau about his rights.

The young boy gained the *height/became/xmuplf* <u>weight</u> in the summer holidays because he spent it watching TV.

The rich aristocrat bought the *tables/unable/lneuma* <u>horses</u> from the famous breeder for his favourite jockey.

The new gun laws helped few *people/column/oahemi* <u>cities</u> in the south of the country to reduce gun crime.

The old athlete played the *sports/showed/neirka* <u>events</u> in the tournament despite his age and still won them.

The successful farmer had needed the *horses/stared/kzmkfe* <u>fields</u> in which to keep his free range chickens.

The tired waiter showed the *groups/oxygen/fikiux* <u>ladies</u> to their tables in the extremely busy restaurant.

The journal editor picked two *errors/wonder/qugewm* <u>papers</u> that he wished to publish in the famous journal.

The clever psychologist showed old *methods/passage/otofewn* studies to his students that were fatally flawed.

When in town, the girl picked the *clothes/disease/lneimlv* friends that she wanted to attend her birthday party.

The charity for the homeless needed the *support/whether/ohiklei* clothes to give to those living on the streets.

The reckless boy had broken his *windows/message/humjuna* fingers after he fell from his bike while doing a stunt.

The increased opening hours caused new *workers/happens/kienmvc* tension in the office between the managers.

The credit crunch had struck the *parents/bedroom/wikuikz* schools that were council funded in central London.

The battling army needed the *support/fingers/oaigimw* weapons if they stood a chance of winning the terrible war.

The couple had chosen the *friends/started/xvksnlz* schools in the area they would consider sending their son to.

The girl was teased for having old *answers/becomes/qeisnkm* parents that were unable to pay for a summer holiday.

The elderly ladies shared old *clothes/contain/ckernmw* stories of their childhood with one another all evening.

The mother was taking the *letters/brought/lnsmoka* friends to the restaurant when her car ran out of petrol.

The hospital DJ played fun *records/clothes/cluvnma* stories on the radio that some of his listeners had written.

The new charity needed its *members/appears/nxmlwui* workers to be on call for emergencies during the night shift.

The business man sought the *records/suppose/kajumai* figures from his bookkeeper before going to see the tax man.

The young couple became new *parents/storage/hxcmxlw* friends after going on holiday together to Switzerland.

The football team wanted the *success/depends/iwqpinh* support from their manager that he had originally promised.

The author hoped her sons would become top *friends/purpose/msokimo* writers when they were older just like her.

The poor employee was giving the *answers/believe/nzmxeio* reasons as to why he was late when he was sacked.

The power outage did not affect the *schools/realise/xkowmlo* streets in the north of the city during the storm.

The social worker listed few *factors/between/neczxmn* reasons as to why the child should go into foster care.

The school inspector closed all *windows/animals/mckuekm* schools in the area that he argued were under performing.

The old receptionist opened all *windows/persons/hskhoao* letters that were addressed to her manager every day.

The oil tycoon was having new *trouble/require/nzmxcie* success in the south of America after striking oil twice.

The kind woman helped her *animals/silence/hwemnla* friends in the village to learn the art of oil painting.

The smart policeman sought any *records/imagine/urmnsnx* reasons as to why he should not prosecute the suspect.

The black plague struck the *workers/getting/olunoki* streets on the outskirts of town where the peasants lived.

The PE teacher taught the *methods/windows/mksnxcv* classes in the local village school where his wife also taught.

The newspaper had to report the *stories/operate/rwmskli* results on the local elections the very next day.

The teacher was giving the *flowers/neither/vfusmna* classes on self defence to help the vulnerable community.

The men were giving the *flowers/operate/oiugmni* weapons to the drug runners who were also ruthless criminals.

The lady had to return the *animals/nuclear/vheltim* clothes to the shops because they were far too small for her.

The dyslexic child was having the *success/include/cnqpiur* support with his writing skills that he deserved.

The Olympian was having big *success/instead/kcmxhti* trouble trying to perfect his high jump after his injury.

The supply teacher taught art *classes/evident/tcmkiaj* history in the past but she hated it as it bored her.

The teenager wanted his *clothes/suppose/hnmclia* freedom from both parents so he moved out and got himself a job.

The lively child had wanted fun *animals/husband/xlianms* stories to be read to her before going to sleep.

The sleeping cat looked too *strange/attempt/uikltih* settled as she slept for her owner to consider waking her.

The head teacher wanted the *history/follows/cmnzoia* science in her class to be more fun for the children.

The helicopter pilot was flying the *workers/failure/rmajiom* weapons to the stranded soldiers as ordered.

The old women picked red *clothes/receive/htmonie* flowers from the beautiful gardens of the grand manor house.

The athletic man had broken two *flowers/explain/cxzoila* records at the Olympic Games despite being ill.

The caretaker of the building locked the *offices/balance/iomlnmu* windows in the factory while he was on duty.

The journalist was asked to report the *tragedy/provide/mfuourw* stories on the devastating conflicts abroad.

The sisters had become big *friends/prevent/euhwila* artists in their adult years, just like their father had.

The unusually strong winds had broken the *flowers/thought/smnlezm* windows in the old derelict farmhouse.

The freezing weather killed the *animals/herself/kfunmia* flowers in the lady's garden much to her disappointment.

The new company tried to obtain the *support/achieve/iklnxmz* offices in the prestigious building in London.

The generous bursary had helped all *schools/version/jcwmnle* parents in the local area to buy essential books.

Appendix R

An example Participant Information Sheet for the eye tracking experiment involving the EyeLink 2000 (Experiment 7).

PARTICIPANT INFORMATION SHEET

INVESTIGATING MISLOCATED FIXATIONS

INVITATION TO TAKE PART IN A RESEARCH STUDY

You are being asked to take part in a research study in which your eye movements will be monitored while reading a collection of short sentences. The data collected will be used for the completion of my PhD, which is supervised by Dr Wayne Murray. This research is funded by the Economic and Social Research Council (ESRC) and the School of Psychology.

PURPOSE OF THE RESEARCH STUDY

The purpose of this research is to look at the way in which people use their eyes and how attention is distributed during reading. You will be required to read the sentences that appear on the screen in front of you, just as you would normally read a book or paper. On a few occasions, you will be required to answer a comprehension question relating to the sentence you have just read. Participation in this research would benefit me as it is a necessary component for the completion of my thesis.

TIME COMMITMENT

It is estimated that your participation in the experiment will take approximately one hour and twenty minutes. You will only be required to attend one session, which will take place in an EyeLink laboratory in the Scrymgeour Building.

COST, REIMBURSEMENT AND COMPENSATION

Your participation in this study is voluntary. You will receive either £5 payment or 4 course credits after completion of the testing.

RISKS

There are no known risks for you in this study.

TERMINATION OF PARTICIPATION

You may decide to stop being a part of the research study at any time without explanation and without penalty. If you withdraw from the study, you will still be paid for your contribution.

CONFIDENTIALITY/ANONYMITY

The data collected do not contain any personal information about you. No one will be able to link the data you provided to your identity and name. The data will be seen only by the researchers and will not be made available to anyone else. Should this research be published, your identity will not be revealed.

FOR FURTHER INFORMATION ABOUT THIS RESEARCH STUDY

I will be glad to answer your questions about this study at any time. You may contact me at email address: ljwakeford@dundee.ac.uk; office number: 01382 388129. If you want to find out about the final results of this study, please feel free to contact me.

The University Research Ethics Committee of the University of Dundee has reviewed and approved this research study.

Appendix S

An example Instruction Sheet for the eye tracking experiment involving the EyeLink 2000 (Experiment 7).

PARTICIPANT INSTRUCTION SHEET

Instructions for Eye Movement Experiment

In this experiment I will monitor your eye movements while you read a number of sentences. Sentences will be presented to you one at a time on a screen. All you need to do is read each sentence normally, as if you were reading it in a book or a newspaper.

Before each sentence appears, a fixation point will appear briefly on the left hand side of the screen. Please look at the white dot within the marker as soon as it appears. After a short delay, the sentence will then appear for you to read. **It is important that you remember to look at the marker before reading each sentence.**

Once you have finished reading each sentence, press the top right hand button and you will then be presented with a series of crosses (xxx xxx), when you are ready for the next sentence, simply press the top right hand button again and the next sentence will appear on the screen.

Following a small random sample of sentences you will be asked to answer a simple question that relates directly to the sentence you have just read (this will appear on the screen instead of the crosses). The required response will either be a yes or no. If the answer is yes, please press the top right hand button, if the answer is no, please press the top left hand button. Once you have answered, please fixate on the fixation point (the white dot within the marker) in preparation for the next sentence.

Periodically, I will ask you to look at a series of fixation points on the screen, again please look at the central white dot within these markers as soon as they appear. This is used for calibrating the apparatus and only takes a few seconds but it is important, so please concentrate. **Also, throughout the experiment is vital that you keep as still as possible while the measurements are being made.** Please keep your head still and try not to move your arms when you press the response buttons (arrange the buttons so that they lie comfortably under the index finger of each hand).

Before you begin the experiment we will go through some practice items to allow you to become familiar with the task. Please use this time to ask any questions you may have and to make yourself comfortable.

Appendix T

An example Consent Form for the eye tracking experiment involving the

EyeLink 2000 (Experiment 7).

CONSENT FORM

INVESTIGATING MISLOCATED FIXATIONS

The purpose of this research is to look at the way in which people use their eyes and how attention is distributed during reading. You will be required to read the sentences that appear on the screen in front of you, just as you would normally read a book or paper. On a few occasions, you will be required to answer a comprehension question relating to the sentence you have just read.

By signing below you are indicating that you have read and understood the Participant Information Sheet and that you agree to take part in this research study.

__________________________ _________________________ ___________

Participant's Name (printed) Participant's Signature Date

Laura Wakeford

__________________________________ ______________________________

Printed name of person obtaining consent

Signature of person obtaining consent

Appendix U

Experimental items for Experiment 4. Italicised words represent the high and low frequency critical nouns, respectively.

The unseasonal snowfall covered *homes/lawns* with a dusting of powdery snow.

The savvy retailer was selling *paper/herbs* at a marked down price.

The online company was selling *items/beads* on Ebay for a large profit.

The forensic scientist removed *blood/mucus* from the cloth for further analysis.

Frederic had a real talent for writing *music/prose* that young people could relate to.

The arrogant doctor ignored *signs/scans* that suggested his patient's cancer was back.

The reckless journalist visited *areas/zones* in the war torn city that were not safe.

The clever detective noticed *glass/resin* on the ground by the body.

The impressive instrument located *water/ozone* on the surface of the planet.

The strange sounds worried *girls/monks* staying in the old monastery.

The successful artist painted *trees/swans* with a great deal of skill.

The stubborn teenager ignored *rules/pleas* from his father and dated the girl.

The devious solicitor altered *facts/deeds* to suit his client's case.

Rod was carefully placing *books/trays* back on the shelf when Audrey entered the room.

During his lunch break, Steve ordered *water/juice* and a panini from the sandwich shop.

The lawyer's words created *doubt/angst* among many of the jury members.

The friends were sharing *looks/grins* of amusement following their tutor's mistake.

The shop reluctantly reduced *books/socks* that were not selling well.

The property developer had removed *money/scrap* from the old bank before the collapse.

The dedicated specialist had studied *blood/coins* for many years indeed.

The researcher had studied *girls/twins* who were born in the same hospital.

The café owner was cutting *costs/salad* in her small yet profitable kitchen.

The builder had exposed *walls/slate* by removing the horrible 1960s tiles.

The friendly students enjoyed sharing *ideas/chips* with one another while at lunch.

My brother's improved grades pleased *father/tutors* who had worried about his future.

Under the terms of the will, the lawyer divided *income/assets* among the man's family.

On entering the room, Mary noticed *coffee/grease* on the new kitchen table.

The corrupt police force allowed *murder/crooks* to go completely and utterly unpunished.

The farmer's son was feeding *cattle/calves* in the field when his mobile phone rang.

Rose had frequently enjoyed *summer/hiking* in the peaceful English countryside.

The new documentary exposed *nature/eagles* as being cruel yet magnificent.

Maria and Neil ordered *dinner/cheese* and wine for everyone at the posh restaurant.

The lazy postgraduates had enjoyed *coffee/gossip* a little bit too much.

The famous scientist located *pieces/chunks* of the meteor in the vast dessert.

The young professional handled *events/leases* for the Royal family last summer.

The elderly gardener allowed *nature/tulips* to grow wildly in the secret garden.

The Olympic Committee were seeking *cities/arenas* large enough to host the games.

The president had offered *action/regret* to the families following the disaster.

The elderly lady offered *dinner/trifle* to her two young granddaughters.

The young boy had decided *school/karate* would be much better now his friend had joined.

The climatic conditions had delayed *winter/hikers* in the mountainous Canadian Rockies.

The talented instructor trained *people/medics* on how to pass the course successfully.

The heroic fireman carried *people/miners* out of the rubble following the explosion.

The tired explorer reached *ground/cliffs* that he recognised as being near to home.

The television presenter visited *cities/fjords* situated on the coast of Norway.

The letter had invited *guests/jurors* to provide feedback and recommendations.

The frustrated mother kept finding *things/towels* that her son had thrown on the floor.

The internal memo advised *anyone/nurses* working within the facility to get vaccinated.

The hotel manager reminded *staff/maids* to be vigilant after a series of thefts.

The small institute had always produced *music/spies* of the very highest standard.

The store manager recorded *sales/debts* in the document for his Area Manager to see.

The art students were painting *women/ducks* sitting beside a big blue lake.

The helpful secretary replaced *paper/toner* in the photocopier when it ran out.

The student was studying *poems/hymns* that had been written many years ago.

The remote location had isolated *women/seals* during the cold winter months.

The cat shelter had assigned *names/cages* to the animals in their care.

The humanitarian charity provided *cover/tents* for people following the disaster.

Robert very carefully examined *glass/flies* that he had found in a box.

The holiday resort employed *staff/chefs* to work in their kitchens.

Claire had been teaching *girls/chess* in the private single-sex school.

The cosmetic surgeon improved *teeth/scars* using a special lightening procedure.

The doctor had frequently referred *cases/peers* to the specialist for further advice.

Jamie had always attended *class/choir* to please his controlling parents.

The special pass entitled *press/media* access to the grand ballroom.

The biologist had compared *cells/worms* from several different locations.

The young lady compared *costs/cakes* from several suppliers before committing.

The receding tide had revealed *areas/caves* on the cliff face Jim had not seen before.

The men were responsible for carrying *cases/sofas* up to the firm's new office space.

Among other things, the surveyor examined *homes/beams* for any evidence of woodworm.

The man was building *walls/sheds* for his client when he became ill.

Amanda had carefully prepared *tests/robes* for the final year students.

The diligent worker prepared *plans/menus* to show his fussy client.

The gun manufacturer equipped *police/squads* with their new powerful firearms.

The traveller had imagined *winter/tribes* to be more hostile on the island.

The old man believed *summer/ninety* was still young, much to his son's amusement.

The sanctuary had released *horses/tigers* back into the wild where they could be free.

As requested, Pippa carefully arranged *events/steaks* for her new important clients.

The bad weather affected *houses/drains* in the centre of the old town.

The large company supplied *energy/jewels* to the new business in London.

The developer had acquired *places/motels* in the north of America to renovate.

The two brothers had always attended *church/mosque* in their younger years.

The loud bang startled *cattle/ponies* that had been grazing peacefully in the field.

The questionnaire had measured *demand/uptake* for the new reporting system.

The beautiful footage inspired *people/divers* all over the world to participate.

The chemical reaction produced *energy/sparks* that no one could have anticipated.

Ed found that therapy often relieved *stress/sorrow* in patients suffering a bereavement.

The creative student was studying *design/sewing* at her evening class in the city.

The girl was watching *groups/swarms* of angry protesters on the street.

The Prime Minister rejected *policy/topics* proposed by the back benchers.

The man with OCD was obsessively cleaning *things/medals* that he had been collecting.

The embarrassed institution provided *guests/alumni* with a full and heartfelt apology.

The large farm supplied *horses/grapes* to the organisers of the fete.

The statistician had realised *growth/scurvy* was on a steady decline.

The maid was paid for cleaning *houses/cabins* at the exclusive holiday resort.

The Formula One driver received *points/cheers* for winning the race in Milan.

The dying man selected *family/nieces* and two friends to whom he would leave his wealth.